Perceptual Content and Perceptual Justification

by

Steven David Yamamoto

B.A. University of California, Berkeley, 2007
M.A. Brown University, 2012

Thesis

Submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in the Department of Philosophy at Brown University

PROVIDENCE, RHODE ISLAND

MAY 2018
© Copyright 2018 by Steven D. Yamamoto
This dissertation by Steven Yamamoto is accepted in its present form by the Department of Philosophy as satisfying the dissertation requirement for the degree of Doctor of Philosophy.

Date _____________  
Christopher Hill, Advisor

Recommended to the Graduate Council

Date _____________  
Brian McLaughlin, Reader

Date _____________  
Adam Pautz, Reader

Date _____________  
William H. Warren, Reader

Approved by the Graduate Council

Date _____________  
Andrew G. Campbell, Dean of the Graduate School
Curriculum Vitae

Steven David Yamamoto attended public schools in Southern California for K-12 education before attending Orange Coast Community College. He completed breadth transfer-program coursework there before transferring to the University of California, Berkeley, where he earned a Bachelor’s degree in Philosophy with High Honors. He then pursued a Ph.D. at Brown University in the Department of Philosophy. He has taught at Brown University (including teaching several times for the Summer@Brown Program), as well as at Bridgewater State University.
Preface and Acknowledgements

I would first like to thank my advisor, Chris Hill, for all he’s done for me during my time at Brown. The dissertation absolutely wouldn’t have been possible without the benefit of his guidance. Of course, I am very fortunate to have had the benefit of Chris’ involvement as a philosopher. Under his advising, the dissertation came together, and in virtue of his input the dissertation was greatly improved. But I was also the beneficiary of years of patient, caring mentorship. I could not have managed the dissertation without all Chris’ thought and action on my behalf, the prompting to take advantages of opportunities that scared me, the long meetings, the encouragement, and the excellent advice. For these things and everything else, I am deeply, deeply grateful.

I would like to express my sincere thanks to my committee members, Brian McLaughlin, Adam Pautz, and Bill Warren, for their thoughtful comments and excellent discussion. They provided me a great foundation for extracting better papers from the dissertation and moving forward with the material. I am also indebted to Susanna Siegel, whom I respect a great deal, who is responsible for my interest in almost everything I discuss in the dissertation, and who provided me valuable draft comments and support. Further, I am indebted to Jaegwon Kim who, besides being a major, positive part of my graduate education, served on my committee prior to retirement, and helped shape the project.

I am extremely grateful to Alex King for being my “dissertation buddy”, for discussion, and for all her support, advice, and care from our early times as struggling, uncomfortable graduate students to the present day. I am grateful to Nic Bommarito for frequently discussing the dissertation with me from the point of view of a different subfield, as well as for years of similar personal support. I’d also like to thank David Bennett, for his helpful comments on an early draft, as well as varied other forms of helpful feedback throughout the process, and for all his encouragement and perspective.

I would also like to thank Brown University for years of financial support and would like to thank Brown’s philosophy department for creating a context in which this dissertation was possible for me to complete. I felt I fit in there, and felt that the collegial, supportive faculty and grad students comprised an ideal environment for my growth and intellectual development. To all the friends I met there, and all the faculty members who make the department such a good place for graduate students, you have my sincere gratitude.

Finally, I want to express my most heartfelt thanks to my spouse, Julia Nedry. Though she didn’t contribute to the content of the dissertation, in all other respects I very strongly feel this is something we’ve done together. It absolutely couldn’t have happened without her. She’s been my friend and partner throughout all of it, and for that, I’m extremely fortunate.
# Table of Contents

Introduction 1

Chapter I
Affordances 11

Chapter II
Rich Content in Visual Experience 43

Chapter III
Cognitive Penetration 84

Closing Remarks 137

Sources 142
Introduction

Most of us share certain intuitions about what our conscious visual experience is like. Imagine you’re flipping through a magazine, and you stop for a moment to gaze at the face of your favorite Hollywood heartthrob. What is that experience like? More specifically what properties are represented in your conscious visual experience? Most will agree that your visual experience represents certain shapes, colors, and spatial properties: your visual experience probably involves, for example, two angled, dark, fuzzy-textured surfaces (some properties of eyebrows), and a protrusion of a certain shape (some properties of a nose). But is that all there is to the visual phenomenology of looking at a face of an arresting beauty? Is your visual experience of his breathtaking visage before you exhausted by things like shapes, colors, and illumination patterns? Some have argued that this is all there is. But, especially recently, many philosophers and empirical researchers have argued that an account on which the content of conscious visual experience is restricted to lines and color amounts to an account that is woefully incomplete. An account that has room only for these basic color and shape properties is unable to account for the way in which, in our visual experience, we don’t just visually experience a bunch of lines and colors in a face-ish shape. We visually experience what we’re seeing as a face – and moreover, as a beautiful face. If this is true, we require a

1 Though, of course, even among those who accept representational perceptual content, there are those who disagree. See Nanay 2012, discussed in Chapter I of the present work.
theory that allows for much richer conscious perceptual content. So, what does figure into the content of conscious visual perception, and how can we investigate the matter?

What follows is an investigation of the contents of conscious perception, conducted in accordance with two main motivations. The first is partly metaphysical and partly empirical. Historically, philosophers and psychologists have been deeply divided concerning questions about lines of demarcation between perceptual and conceptual content. In philosophy, for example, there has been a long controversy about whether awareness of causal relations is perceptual or instead depends on conceptually grounded interpretations of perceptual experience. Hume, perhaps the best-known historical entrant in this debate, famously argued that one cannot directly perceive “necessary connection”, and must instead make a (questionable) inference from observed sequences of events to a causal relationship. Psychologists have quarreled about these issues as well, especially since the work of Michotte in the 1940s. And there have been many controversies about boundaries between perceptual content and conceptual content in other areas, including controversy about whether such a boundary is theoretically useful at all (See Gibson 1979). These controversies are very much alive at present, and they have been joined by newer ones of the same general type. For example, there is now the issue, debated both by philosophers and psychologists, of whether affordance properties like climbable and graspable are “directly perceived” (in a sense championed by J.J.Gibson), or if there is, instead, still good sense to be made of saying that these affordance properties are registered at the conceptual level, but not as part of one’s perceptual content.

A further reason for interest in perceptual content, noted by Siegel (2010), is that our best theory of the content of perceptual experience may constrain programs of
empirical research. For example, if we wind up with good reason to think that the content of conscious visual experience is thin, researchers looking into the neural correlates of conscious visual experience should expect to find these correlates in relatively early areas visual information processing.

It has been a goal of my research to contribute to resolving some of these disputes, and I think I have had some success in this endeavor. In what follows, I defend the conceptualists; that is, I argue that the view that conscious perceptual experience is exhausted by low-level properties is a well-motivated, viable view consistent with the empirical evidence. I argue for this view in Chapters I and II. In general, I argue that experiments that have been taken to support claims of thick perceptual content are actually perfectly well-understood as indicating thin perceptual content that is elaborated considerably by conceptual interpretations.

The importance of investigating the content of perception is not limited to metaphysics of perception and philosophy of cognitive science, however. My other main motivation in this examination of perceptual content is epistemological. In thinking about what kinds of things we can know and what kinds of things we might be justified in believing about the world, perceptual experience holds a place of particular importance on most accounts. At least intuitively, it’s difficult to deny that perception plays a special epistemic role, as we suggest when we say things like “seeing is believing” or when we express the remarkable dissonance we feel when we “can’t believe what we’re seeing”, and much philosophical thought has been devoted to analyzing and accounting for the special epistemic power perceptual experience seems to have. Foundationalist theories of justification, for example, often take perceptual experience to be an “unjustified justifier”.
On these theories, perceptual input forms an indispensable part of the “foundation” of basic beliefs. Thanks to perception, the foundation is broad enough to support the host of other justified beliefs about the world most of us take it a good theory should account for. Even coherentist theories of justification sometimes use the likely inclusion of beliefs like “My perceptual experience is reliable” in the coherent belief sets of most human beings to ward off complaints of epistemic disconnection with the world. On almost all accounts, what we can see is intimately linked with what we can know.

The present work is focused on exploring several issues surrounding the content of conscious perceptual experience itself, as well as the role of that experience in justifying belief. It’s an exciting time in the long and storied history of the intersection between epistemology and philosophy of perception. Recent attention to the “richness” of conscious visual experience, and questions about the modularity of perception have opened the door to a number of questions about what experience can do for us epistemically, and the present work is dedicated to an exploration of several of these questions.

The metaphysical/empirical issues that will be discussed have very direct application to epistemological issues. On one popular type of foundationalist view, for example, we often have prima facie justification for believing things are as they perceptually appear to us. Determining the content of perceptual experience, then, should be a central concern of any such foundationalist attempt to determine what kind of epistemic contact we can have with the world. Richer perceptual content would mean more complex beliefs would have this direct prima facie justification. Can our perceptual experience directly justify our beliefs about “rich” properties like the anger of a
companion or the hazardousness of one’s environment? Or can it, at best, directly justify only “thin” perceptual beliefs about things like color, shape, illumination, and motion?

For example, if perceptual content is limited to these low-level properties, then my visual experience upon looking at an apple can represent only properties like redness, spatial relations, and certain surface texture patterns. If I’m an epistemically responsible agent with no defeaters, I then form perceptual beliefs about shape, color, and spatial relations. I might, then, together with some conceptual resources, infer that there’s an apple on the table in front of me, and then infer that what’s in front of me is edible. When we subject these beliefs to a foundationalist evaluation, we might say that my beliefs about the shapes and colors in front of me have direct prima facie justification; that is, they are prima facie justified just in virtue of their connection with my visual experience presenting those properties. My belief that what’s before me is edible, on the other hand, is at best mediately justified. I made good inferences, using background information, from the justified beliefs about the shapes and colors in front of me to a justified belief that what’s in front of me is edible. However, if perceptual experience represents not only low-level properties, but higher-level properties as well, then it might be, for example, that I don’t merely have a color/shape experience followed by cognitive operations which produce beliefs about an apple, and then edibility. Instead, I just have a visual experience as of an apple, or as of an edible thing. Such a view is not hard to elicit support for – it is supported by a fairly robust program of empirical research, as well as some particularly tempting demonstrations and examples in the literature (imagine looking at a page of a written language entirely foreign to you, and then imagine the experience of viewing that same page after becoming highly proficient with that language). If perceptual content
involves these high-level properties, inference need not enter into the question of whether or not my belief that there is an edible thing in front of me is justified – edible-ness in front of me was just part of my visual experience, and that experience directly justifies my belief. In this way, the more we stretch the boundaries of the content of perceptual experience, the more directly justified beliefs we can have, and, as some rich content theorists advertise, this can be a way in which rich content stands to bear on a number of hotly disputed issues.

For example, if visual experience is rich, the foregoing considerations offer the promise of new ways of understanding the kind of epistemic contact we have with sociocultural properties and properties we take facial expressions to express. It may be the case, for example, that we do not just experience the lines, shapes, and movements of a face and then draw upon conceptual resources to infer that the person in front of us is in terrible pain – we simply have a visual experience as of a pained person. Similarly, justification for the attribution of aesthetic properties, for example, might well be affected in this way: you simply, as suggested in the case mentioned earlier, see the beauty of the face before you (see, for example, Stokes (forthcoming)).

One can see rich content theory, if true, potentially having a similar impact on any area in which we have had some difficulty seeing how one could justify an inference from the low-level properties of the world to a high-level property: when combined with certain foundationalist sympathies, rich content makes the problematic inference cease to be a part of the justification question. The debate about how we acquire knowledge about other minds, for example, mainly dominated by discussions of theory-theory and simulation theory, might have a new competitor in “direct social perception”, according
to which, for example, you might just directly visually experience your partner’s
disapproval of what you’ve done (See Krueger (forthcoming) for such an account). Or
perhaps, contra Hume, we can be justified in believing one event causes another after all,
since perhaps we do directly experience the causal relationship over and above the
repeated witnessing of events of one type following events of the other.

How thick the foundation for our perceptual beliefs might be is one of the central
concerns of this work. I investigate it by investigating the nature of perceptual content.
Do we have good reason to accept a rich view of perceptual content? In the first two
chapters, I criticize some general methods for evaluating rich content theory
phenomenologically, and then proceed to examine several of the best-developed and best-
empirically-supported individual cases, as well as some related philosophical issues
attaching to these individual cases. This includes special attention to perception of
“affordances” in its own chapter, due to affordance perception being particularly well-
developed, empirically and philosophically. In this chapter, I also explore the Gibsonian
idea that perception is fundamentally affordance-based, as well as the worry that our
affordance-oriented evolution gives us reason to think we’re never in perceptual contact
with observer-independent properties of the world.

Ultimately, I argue that we do not, in general, have good reason to accept the rich
content view, and outline a general strategy of argument against empirical work
purporting to challenge the thin view.

Additionally, while a primary focus in the first two chapters is on what is in the
content of perceptual experience and on what kind of beliefs might count as directly
justified as a result, the third chapter will explore how experiences come to have the content they do, and what the epistemic significance of that might be.

In the first two chapters, I consider the epistemological ramifications of “rich” and “thin” theories of content for views according to which perceptual experience holds a special epistemic status. The strengths of such views themselves, however, have come under new discussion in relation to “cognitive penetration” cases. These cases, in which the character of our perceptual experience is influenced by our higher-level states like beliefs and desires, have brought potentially new angles of discussion to the fore – particularly since there seems to be empirical support for the claim that such “cognitive penetration” occurs fairly regularly. For example, in one oft-cited experiment, subjects tended to match a monochromatic line-drawing of an African-American to a darker comparison patch than when matching a monochromatic line-drawing of a Caucasian-American (Levin & Banaji 2006).

Roughly, the question I investigate here is how we should think about experience’s ability to provide justification in light of cases like these. If I believe a man looks dangerous on the basis of a “rich” visual experience of him as a dangerous man, then on the kind of internalist, foundationalist picture described earlier, I should thereby be directly justified in believing him to be dangerous. However, something seems wrong with this picture if I only have an experience as of a dangerous man because I antecedently believe the man to be dangerous for no good reason at all. In this case, there’s clear intuitive force to the claim that such a belief isn’t justified, which threatens to undermine the sort of perceptual foundationist theory mentioned earlier (and also threatens to cause problems for beliefs about the reliability of experience that coherentists
are wont to include in the sets of beliefs they regard as justified). The question under
discussion, then, is what cases like this might mean with respect to the epistemic role
perceptual experience is best thought of as playing. This is the other main topic under
discussion in this work, and the primary subject of the third chapter.

In the third chapter, I examine a number of accounts of the proper epistemic role
of experience in light of cognitive penetration cases. I accept that cognitive penetration
generates intuitions that can be, in a certain sense, unfriendly to the sort of foundationalist
view on which we are directly prima facie justified in believing things are the way they
appear to us in visual experience. However, I defend this sort of view from an over-
reaching attack based on those intuitions. In the end, I argue that the profile of intuitions
the cognitive penetration cases elicit can be handled well by a pluralistic account of
justification combining aspects of phenomenal conservatism and reliabilism. I defend this
account against several rival “mentalist” proposals – proposals which include a non-
reliabilist account of justification which, nonetheless, is supposed to respect the same
intuitions reliabilism does. According to these “mentalist” views, justification need not be
a matter of factors the subject has access to, but justification is still a matter of one’s
mental states (and relations among them). I argue against these mentalist competitors’
claims to be the most promising account in light of considerations raised by cognitive
penetration cases. Particular focus is given to the account presented in Susanna Siegel’s
recent *The Rationality of Perception*, in which she claims that the perceptual experiences
*themselves* are the product of processes of inference (and sometimes bad inference, as in
the kind of troubling cognitive penetration case above), and should themselves be
rationally evaluable and redound on the agent’s rational standing.
Chapter I: Affordances

Here are some of the ways in which we engage in successful visually-guided action every day of our lives. We successfully grasp objects in the right places and in the right ways with our hands so as to lift and manipulate the objects. We capably walk down obstacle-laden pathways, avoiding fallen branches and crossing uneven pavement. We walk through too-narrow doorways and up and down staircases without bumping ourselves (very often). We throw things, we catch things, we drive cars and high-five people, and all to a pretty astounding degree of success - as far as the activities above go, we're pros.

How do we so smashingly successfully coordinate our visual information and our actions? The traditional answer is something like the following: we use the impinging of light on our retinas to create, through a series of computational transformations, a perceptual representation of the spatial layout around us. This representation amounts to our system’s “best guess” given the incomplete information our sensors provide. We then use this representation to plan out our behavioral engagements with the environment - our behavioral success is due to a good match between the model and the layout of the outside world, at least insofar as it makes a difference to our actions. This picture, however, famously generates a number of difficult puzzles. For example, one might wonder what explains the match between the representation and the world. Further, is it a consequence of this kind of picture that the viewer is problematically epistemically
"locked-in", in the sense that she has access only to a representation of the world that is (to a philosophically troubling extent) of her own creation?

In the mid-to-late 1900s, James J. Gibson, dissatisfied with the traditional story above, developed a new account of perception and action, suggesting that the assumptions that came with the traditional model-inside-the-head sort of account doom it to failure. He argued that we'd gone fundamentally wrong in theorizing about the organism and the environment separately, looking for a way to get one into the head of the other to explain action. Instead, he argued, we should develop an understanding of perception on which we ought to be understood not as fundamentally perceiving properties of the environment (like lengths and heights) for the sake of generating a mental model of the viewer-independent world, but rather as directly perceiving the viewer's opportunities for action in themselves, tying perception and action together in a much closer way: you successfully grasped the cup because perception just involves perceiving the potential action the cup afforded - you saw the graspability.

In the present chapter, we'll look in greater depth at Gibson's account and its departure from the orthodoxy in order to see what we might learn about perceptual content generally, as the Gibsonian view has far-reaching implications for a philosophical investigation into perceptual content. Notice, for example, that while Gibsonians are concerned to provide an account of perception on which we aren't, so to speak, trapped in a perceptual world of our own making, what we do wind up in direct perceptual contact with on the Gibsonian proposal is viewer-dependent properties. How happy we ought to be with this as a way of being in perceptual contact with "the outside world" is an issue
we'll take up further on in the chapter in discussion of a similar point made by Kathleen Akins.

Overall, the main focus of the chapter will be the Gibsonian claim that we are in perceptual contact with these opportunities for action, or "affordances". There are three separable claims in Gibson's writing, each of which I wish to comment on in what follows. First, there is the plausible claim that perception is at least sometimes of affordances. This claim would be false if affordances never figured into perception proper (if, for example, it were the case that affordance awareness was always the result of the perception of the viewer-independent Euclidean geometry of the environment plus a very late, cognitive judgment).

The second claim is that affordances are uniquely specified by the light, and so do not require the kind of elaboration and filling-out transformations that the traditional accounts require. That is, there is a one-to-one mapping between whatever affordances we see, and some structural property of the ambient light array (for each, under normal conditions).

Finally, there is the claim that affordance perception is the most basic form of perception, and perception of traditional properties is somehow derived from this basic affordance perception. The last of these is of particular importance for thinking about perceptual content generally. For one thing, if perception is, in fact, fundamentally a matter of contact with affordances (say, if its most basic representations are representations as of affordances), this will be a tremendous change to the orthodoxy of the perceptual sciences. This would also have wide-reaching theoretical ripples, reaching into areas like evolutionary theory (since the basic unit of perception will become what is
advantageous and disadvantageous to the organism), and action theory (since perception
would be recast as fundamentally action-oriented).

Further, if the case can be made that our perceptual systems are primarily in the
business of guiding behavior rather than producing anything like a map of the organism-
independent world, this will be grounds for eliminating the notion of representation
altogether. Two important such lines of thought will be examined further on in this
chapter.

Finally, if perception fundamentally represents affordance properties, this bears
on theories about conscious perceptual content. We might wonder what kinds of
properties figure not only into vision generally, but into our conscious visual
phenomenology. Some have thought visual phenomenology is "thin" in that we can only
have a conscious visual experience as of what might be called "low level" properties like
sizes and shapes and colors. On the other hand, some have argued that visual
phenomenology is "rich", in that we can have conscious visual experience as of the
graspability of that pen, or the semantic properties of that string of symbols, or the natural
kinds instantiated before me (the pine-tree-ness of those pine trees). If it is affordance
properties that are the basic representata of our perceptual systems, this would make
implausible what is, by far, the most popular thin content account of affordance-related
perceptual content: that visual experience is as of sizes and shapes and colors and so
forth, and it is at a later stage of processing (some stage too late in processing to be part
of our conscious visual awareness) that affordance properties appear. This line will be
examined at the end of this chapter.
The claim that affordance representation is fundamental is the one I will devote the most attention to, and will criticize at greatest length. I will also discuss the aforementioned philosophical worries about representation in general which are related to this claim, and I will argue that rather than supporting this claim, empirical work in the area has provided an increasingly plausible alternative to it, as well as a tool to resist similar arguments from a more a priori angle. Further, I argue that there is some positive reason to think, to the contrary, that perception of traditional properties like size and shape is more basic than perception of affordance properties. Last, I intend to dispute the claim that affordance properties feature in visual experience independently of our experience of traditional properties - that is, I argue against the fundamentality of affordance representation in the realm of conscious visual experience.

The Gibsonian Tradition

Affordances are the keystone of Gibson's proposed revolution in vision science. Gibson took his views on vision science to be quite radical - as mentioned earlier, he felt the orthodox view suffered a number of deep problems stemming from confusions which he was determined to avoid. One of the most important of these problems Gibson put in the following way in his first book:

"If the mind constructs the world for itself, why does it agree so well with the environment in which we actually move and get about? If space perception is a subjective process, then why are we so seldom misled by illusory perceptions?"

(Gibson 1950)

The dominant line in vision science (which remains the dominant line today) was what Gibson sometimes called the "traditional" or "Helmholtzian" account. According to this
account, vision begins with a time-slice of sensory stimulation insufficient for determining the outside environment. Consider famous examples of underspecifying visual stimuli that give rise to perceptual illusions like the Ames Room illusion - the fixed-point image of the Ames room is ambiguous between an ordinary room made up of mostly right-angles and a bizarrely slanted room. This stimulation is then added to by a subsequent series of computations or, if you like, "unconscious inferences" (a term often attributed to Hermann von Helmholtz) to produce a disambiguated representation of the world, and it is the content of this processed representation to which we have perceptual access. This is why the aforementioned Ames Room produces an illusion: your perceptual system makes sensible (but, in this case, inaccurate) enrichments to the representation in accordance with the assumption that the room is made up of right angles. This is the sense in which "the mind constructs the world for itself": in the words of V. S. Ramachandran, a modern holder of the traditional view, "we are hallucinating pretty much all the time". Gibson's challenge, then, is that the traditional theorist must explain how, if that is the case, it could possibly be that we are, moreover, veridically hallucinating pretty much all of the time.

Gibson and his allies doubt whether this challenge can be answered. Traditional views, they say, take a loan out on intelligence insofar as stimulation underspecifies the environment, but processed stimulation is, in fact, largely veridical or at least does, in fact, successfully guide behavior - recall that we'd like an account of perception to be able to explain the sort of rampant success with visually-guided action mentioned at the start of this chapter. Where is the accuracy coming from? The organism, it seems, must have something like prior knowledge of the outside environment, for otherwise, the organism's
computations would be useless (for example, the Ames room "enrichment" is sensible because, in the actual world, right-angle rooms are very common, and Ames rooms are vanishingly rare). Gibsonians worry that this loan on intelligence can only be repaid either by showing that the prior knowledge is acquired or innate. If it is acquired, a regress looms: the answer to "how does the organism acquire that prior knowledge" had better not be "perceptually", since, on this account, that would be to invoke further unexplained prior knowledge. Gibsonians have similar misgivings about explaining the prior knowledge as innate. This only pushes the problem back a step, as the theorist must now explain why the organism has this innate knowledge by explaining why its ancestors had such knowledge and how it passed on. As Gibson put it:

"The error lies, it seems to me, in assuming that either innate ideas or acquired ideas must be applied to bare sensory inputs for perceiving to occur. The fallacy is to assume that because inputs convey no knowledge they can somehow be made to yield knowledge by "processing" them. Knowledge of the world must come from somewhere; the debate is over whether it comes from stored knowledge, from innate knowledge, or from reason. But all three doctrines beg the question. Knowledge of the world cannot be explained by supposing that knowledge of the world already exists. All forms of cognitive processing imply cognition so as to explain cognition." (Gibson 1979)

Gibson is, here, overstating a bit. It's far from clear that the dilemma holders of the traditional view are in is in principle insurmountable or results from some sort of logical mistake. For example, shifting the problem back to the ontogenetic level by invoking innate knowledge doesn't leave one with the exact same explanatory resources one had at the problem's original level (this is what often makes accusations of "pushing the problem back" so biting). It's certainly not clear that there's not a coherent, plausible position to be had here - in fact modern nativism about certain concepts is a coherent, plausible position (Carey 2011). But for our purposes, we need not hold Gibson to anything as strong as the
above sounds. It will suit our purposes merely to note, with Gibson, that at least the traditional view takes on a heavy explanatory burden by taking out a loan on intelligence - a burden that we would do well to avoid if possible. This is more than enough to see how Gibson's work and the work of his followers is well-motivated.

**Affordances**

Gibson's new approach to perceptual psychology, what came to be called "Ecological Psychology", was aimed at avoiding the above burdens by avoiding what Gibson saw as the most insidious of the traditional view's assumptions. His valuable insights were many and varied, but of particular interest for our purposes here is the Gibsonian notion of an "affordance" property.

Gibson (1979) introduced "affordance properties" as the term for something like an action opportunity. Further, again, he claimed that the perception of these action-oriented properties is in some sense prior to or more basic than the perception of traditional properties like size, shape, and color, writing: "[W]hat we perceive when we look at objects are their affordances, not their qualities" and continuing further:

"[p]henomenal objects are *not* built up of qualities; it is the other way around. The affordance of an object is what the infant begins by noticing. The meaning is observed before the substance and surface, the color and form, are seen as such."

(ibid p.134)

Gibson was notoriously (and likely intentionally) vague about what, exactly, an affordance ought to be taken to be. As a result, what affordances are varies from account to account, but most take as the starting point Gibson's oft-quoted line, "[t]he affordances of the environment are what it *offers* the animal, what it *provides or furnishes* either for
good or for ill." (Gibson 1979 p.127). Affordances are something like opportunities in the environment for an organism to interact with that environment in some way relevant to that organism. Gibson himself offers a very wide range of affordance examples at different levels, talking in the same chapter both of rigid, non-slippery surfaces affording locomotion and support for animals of certain sizes and configurations (plausibly explainable just in terms of facts about things like surfaces and musculature), but also of mailboxes affording letter-posting and politicians affording deception (certainly not explainable just in terms of facts about things like surfaces and musculature).

Gibson saw affordance perception as a way of seeing "meanings", which turned out to have been in the environment all along, writing "[A]n affordance cuts across the dichotomy of subjective-objective and helps us to understand its inadequacy. It is equally a fact of the environment and a fact of behavior. It is both physical and psychical, yet neither." (ibid p.141).

The letter of what Gibson wrote is (again, likely intentionally) provocative and sometimes difficult to parse, and it is indeed unclear what he meant at times. It is, at least, however, not difficult to read Gibson charitably in outline. Warren (2005) is an excellent example of such a reading of Gibson. He writes that one of Gibson's main valuable insights was that:

"Perception is... pragmatic and task-specific. It need not "solve" general problems such as the recovery of absolute space and time, but must provide contact with behaviorally relevant features and properties such as reachable prey and looming threats."

Notice the way in which affordances are a key notion for a successful Gibsonian theory: Gibson thinks that one gets into intractable difficulties once one assumes that the human perceptual system is in the business of enriching informationally impoverished
stimulus information. One way to avoid this assumption is to make the case that the information *out there in the world* is, by itself, rich enough to explain the content of our perceptual states without requiring the above sort of Helmholtzian "unconscious inference". That is, in couching his theory in terms of "affordances", Gibson is attempting to recast the organism's perceptual states and the environment it's perceiving so that they're close enough that the organism need not bring any "prior knowledge" to the perceptual process.

Whether or not Gibson is successful in this is, however, to a great extent an empirical matter, and Gibson was aware of this. It is the project of ecological psychology to determine what affordances there are for an organism, and then to see if we can figure out how (and whether or not) these affordances might structure the light such that the light might *specify* the affordances, and also how (and whether or not) organisms' perceptual systems actually take advantage of this information.

There is at least some reason to think the Gibsonian project might stand a better chance than the Helmholtzian one in this respect. If Gibson is right and it is the case that positing enriched intermediary representations results in insurmountable problems for a theory, it's likely the case that a “traditional” view on which we represent the full-blooded, Euclidean geometry of the world will come up short. That much information may just not be available in a moment’s worth of light\(^2\). As an alternative to this, it's at least reasonable to think that our search for an account of perceptual content ought to proceed upon a consideration of what it is the perceptual system evolved to do in the first place. That is, recognizing that "[p]erception did not evolve to be the 'passive' recipient of

\(^{2}\) Of course, a modern version of the traditional theory need not be restricted to a single moment. I omit these considerations, as my main aim is to provide a theoretical contrast class.
sense impressions from an 'outer' world, but rather to guide action *in* the world" (Warren 2005) and recognizing that we, in fact, are massively successful in our visually-guided behavior, together give us some reason to think it might be the behavior-relevant features of the environment that we are directly picking up on.

Now, one reason to worry about Gibson's affordance-based understanding of perception in general is that there appear to be counterexamples, or at least terribly difficult cases. My perception of color, for example, seems very hard to try to ground in a more fundamental perception of affordance properties. Does it seem worth seeking out some set of affordances which all and only red things offer in any plausible environmental niche? The existence of such a set seems rather doubtful from the armchair, and I know of nowhere that color properties are discussed in the Gibsonian literature. However, the Gibsonian approach might still be the right approach to handle most or many cases, color aside, and so even if the prospects for grounding color perception in affordance perception seem dim, we might still benefit from restricting the domain of Gibson's claim a bit. It would still be quite astounding (and the resultant philosophical issues discussed below would still concern us) if, color aside, it turns out that, say, we are in basic perceptual contact with none of the basic geometry of the world and, instead, our perceptual representation of these properties is grounded in affordance property perception.

Despite it being clear how to understand Gibson on affordances properties in broad strokes, as accounts vary broadly, having a particular account on the table will be helpful. In what follows, I'll assume Michael Turvey's particularly well-accepted development of Gibson's affordances (Turvey, 1992) on which affordances are
dispositional properties of the environment, and are therefore modal properties. For example, surfaces of a certain size, shape, height, and rigidity (we know them as "chairs") afford sitting for an average adult human in that if such a human were to position herself on the chair in the right way, it would support her in a certain comfortable way. Notice, the agent plays an ineliminable part in picking out the affordance: the ordinary chair affords sitting only for our average adult human with average musculature and coordination and so forth.

So, in sum, I will take the "Gibsonian" view to be the view that we are primarily in perceptual contact with the dispositional affordance properties of things (the sit-on-ability and grasp-ability of the chair), and our perception of properties like the shape and surface of the chair is somehow parasitic on the perception of the affordances.

The chief rival to Turvey's dispositional account of affordances is what is sometimes called the "relational account" of affordances, prominently defended by Anthony Chemero (Chemero 2003). Chemero\(^3\) presents an account on which an affordance is a relation, which he symbolizes as "Affords-ϕ (environment, organism)" which, put more roughly, is something like "The environmental feature affords ϕ-ing to the organism". He spends most of this paper intelligently discussing the relata of this relation, but very little time discussing what the relation itself is, or in what ways, on his view, the relation is non-identical to the modal relation proposed by Turvey. Absent such an account, the most charitable notion to adopt seems to be the generally well-accepted Turvey dispositional account.

\(^3\) The motivations he cites are clever and interesting, reaching back to Strawsonian philosophy of language, but in the present context they aren't clearly more independently plausible than the motivations for a dispositional account, and would be something of an aside, so I won't discuss them here.
Gibson and Fodor and Pylyshyn: Ecological Psychology is a Non-starter

Fodor and Pylyshyn, in a famous, lengthy response to Gibson (1979), claim Gibson’s account suffers from what they call the "trivialization" problem. They argue that Gibson’s claim that "all perception is the direct pickup of invariants" (that is, our perceptual systems "resonate" in some way to the full specification of the affordances in the ambient light array) is vacuously true unless "invariant" is suitably constrained. They write:

"Suppose that under certain circumstances people can correctly perceive that some of the things in their environment are of the type P. Since you cannot correctly perceive that something is P unless the thing is P, it will always be trivially true that the things that can be perceived to be P share an invariant property: namely, being P. And since, according to Gibson, what people do in perceiving is directly pick up an appropriate invariant, the following pseudoexplanation of any perceptual achievement is always available: to perceive that something is P is to pick up the (invariant) property P which things of that kind have." (Fodor and Pylyshyn 1981)

The authors then go on to present an example: perceiving a shoe by virtue of direct pickup of the invariant property “being-a-shoe”, which all shoes have. The rest of the paper is, for the most part, an attempt to try out various ways of constraining "invariants" so as to exclude shoe-perception cases, only to find that each of these either won't work, or would be antithetical to Gibson's anti-traditional-view project.

The complaint Fodor and Pylyshyn make certainly has a valuable thrust behind it, but in the end, seems to be a category mistake. It seems as though Fodor and Pylyshyn might have the wrong idea about what is meant by "invariant" in the Gibsonian literature - they assume "invariants" are properties of objects themselves or of the environment itself. Gibsonian invariants, however, are meant to be structural properties of the ambient light. The idea is that the environmental properties, the affordances, are such that for any visible
affordance there will be some corresponding structural property of the light caused only by that affordance in the ecological niche of the organism it is visible to. The invariant is that structural property of the light, and not any property of the environment itself (such as shoe-ness). Ecological vision scientists investigate whether or not these physical layouts of the environment structure the light in such a way that the light has that structure which remains invariant across changes in illumination, or changes due to the movement of the observer or the environment.

It isn't true, then, that there are no constraints on invariants - they must at least be properties of the structure of the light, and it is certainly non-trivial that they must uniquely be caused only by a certain affordance property within an ecological niche. In fact, the lack of enough such invariants to explain the affordances with which we think each organism is in visual contact should be one of an ecological psychologists' chief concerns (and often it is - some ecological psychologists relax the picture to allow for a one-to-many mapping of invariants to affordances at the cost of some of the theoretical advantages Gibson sought). Perhaps the constraints on what counts as an invariant aren't quite as firm as we might like (again, this seems intentional - Gibson gestured quite widely at the possibility of acoustic, mechanical, and chemical invariants that might await discovery), but they are adequate to disallow cases like Fodor and Pylyshyn's trivial version of seeing shoe-ness. Since shoes vary quite wildly in their surface arrangements, it's likely that there's just no invariant in the ambient array specifying shoeness, as shoeness is, perhaps, too ill-behaved a category (with respect to its surfaces) to produce patterns in the light specifying only and all shoes.

This is not, of course, to say that Gibson's view is not at all wanting in terms of well-defined constraints. A pattern in light might specify more than one thing: recall the
classic, "black dots versus flies". Which of these should a Gibsonian say that I perceive? (This issue will be touched upon later in discussion of Akins (1996).) The point here is only that Fodor and Pylyshyn haven't given us reason to think the view is a non-starter.

Of course, it might turn out that, surprisingly, there is some very high order invariant that a brilliant, mathematically minded ecological psychologist will find that specifies all and only shoes (based in the musculature of human feet, perhaps). But such a discovery would be just that - a discovery. It's not just trivially the case that we can see the property "being-a-shoe" - the specifying information in the light, if it exists, would need to be discovered.

**Threatening Perceptual Representation Generally**

In order to see how the affordance-orientation of our perceptual systems might threaten perceptual representation generally, we can consider Kathleen Akins’ well-known discussion of this potential threat (Akins 1996).

Akins details the way in which many have pursued what one might call the "naturalist" strategy in trying to explain the nature of representation in general or, as she puts it, the "aboutness" relation: First, take (visual) perceptual representation as a simple, paradigm case of "aboutness". Next, identify a "natural" relation which will suffice to explain the way in which our perceptual states are "about" the world; that is, find a "natural" relation between, say, the properties of a basketball and the perceptual states I have when looking at the basketball, in virtue of which my perceptual states can represent its round surface and orange color "out there" (for example, one popular
candidate is the obtaining of a certain causal relationship between the two). Finally, generalize from the perceptual case to mental states generally. My thoughts’ being about my spouse can be explained, the theorist hopes, in the same way my perceptual states’ being about the basketball can.

However, Akins and allies see reason to worry about the "naturalist" strategy. Akins makes the case that our sensory systems don’t need to serve as devices for putting us in touch with the features of the world in anything like the way above. There is no need, she argues, for them to act as faithful reporters of the way the world is. Instead, they just need to guide our behavior successfully - to keep us safe, move us toward food, away from extreme temperatures, etc. In light of this, one might worry that the "naturalist" strategy is misguided: perhaps we shouldn't be taking visual perceptual states as a paradigm case of "aboutness", since it might be the case that our visual system isn't in the business of representing the outside world at all.

One might see the similarities between the concerns mentioned here and the trajectory of Gibson's writing, especially in Gibson (1979). Recall that the right way to think of perceptual systems, on Gibson’s late account, is to think of them as putting us in "direct" contact with affordance properties, or opportunities for action, which are dispositional properties of the environment (Turvey 1981). Much research in ecological psychology is aimed at showing something like this is true on a case-by-case basis. Chairs afford sitting for an organism with the right musculature, food affords nourishment, and it is these properties with which our sensory systems must put us in touch. But, here looms the worry Akins is pointing out: one might think that, if ecological psychologists are right, the properties which sensory systems put us in touch with are too
deeply tied to the needs of the organism to helpfully model the way my thoughts are
about the outside world.

In the remainder of this chapter, I defend the "naturalist" strategy and suggest
reason to think the Gibsonian priority claim is false. I argue first that what work in
ecological psychology has shown is not that we must be in bottom-level perceptual
contact with anything like modal affordance properties, but that it might be that we are in
perceptual contact with more "traditional" visible properties in a way that is very
congenial to our awareness of affordance properties. I will then argue that this result
militates against worries about the "naturalist" strategy brought by Akins by showing on
a case-by-case basis that visual behavioral guidance is a matter of first being in visual
contact with "traditional" visible properties.

**Sensory Narcissism**

According to Akins, for sensory systems to serve as a model of "aboutness"
congenial to the "naturalist" strategy mentioned above, they must be systems which,
“capture the structure of a domain of external properties [and] tell the brain, without
exaggeration or omission, 'what is where.’” (Akins 1996). She invites us to imagine a
fantasy picture of thermoreception in which one’s skin receptors are finely tuned to the
temperature of the surrounding area, producing some sort of clear signal that carries
information about skin temperature. Intuitively, such a picture seems ripe with possible
candidates for a "natural" relation explaining the way in which such a sensory system
would produce states that are "about" skin temperature.
However, our sensory systems appear not to operate in the sense that Akins argues the "naturalist" strategy requires. Instead, they are, as Akins puts it, “narcissistic”: they do not faithfully report on anything in the world, but rather are just concerned with the behavioral success of their possessor.

Thermoreception in actuality is quite unlike the above fantasy picture. For one, thermoreceptor cell activity levels map one-to-many onto skin temperatures - most levels of activity (the states a “naturalist” might hope to take as perceptual states) don’t specify a single skin temperature. Further, thermoreceptors respond "hysterically" to sharp rises and falls in temperature, spiking wildly before settling to a stable level of activity after each sharp change. Further, the magnitude of this spike is dependent not only on the sharpness of the rise or drop in temperature, but also on how extreme the temperature of your skin already is. In light of the above, it’s hard to see the why the thermoreceptive system should be taken to be a temperature detector rather than a skin protector. After all, it delivers an intense warning signal when a dangerous change in temperature occurs, which is all the more urgent if skin temperature is already at a dangerously high/low level.

Akins worries that the thermoreceptive system seems not to be doing anything like what the naturalist originally wanted it to serve as a simple model for. If our thermoreceptive system is supposed to generate representations of the objective temperature states of the skin, it looks to be doing a terrible job. However, if the thermoreceptors are just supposed to prevent certain kinds of avoidable damage to the skin, they’re doing a great job. The larger worry, then, is that these considerations
generalize to all our sensory systems. A sensory system, as Akins puts it, “is not asking, ‘What is it like out there?’ … rather, it is doing something.”

The Fundamentality of Affordance Perception

Recall that Gibson (1979) claimed that perception of what he called "affordance properties" is in some sense prior to or more basic than the perception of traditional properties, writing: "[W]hat we perceive when we look at objects are their affordances, not their qualities". Recall that, on the Turvey account, affordances are modal properties. For example, surfaces of a certain size, shape, height, and rigidity afford sitting for an average adult human in that if such a human were to position herself on the surface in the right way, it would support her in a certain comfortable way. So, as Gibson has it, we are primarily in perceptual contact with the dispositional affordance properties of things (the sit-on-ability and grasp-ability of the chair), and our perception of properties like the shape and surface of the chair is somehow parasitic on the perception of the affordances.

A main source of potential support for this view is the work of other ecological scientists emerging shortly after Gibson’s (1979). Warren (1984), for example, found that the perceptual judgments of subjects who were asked to tell whether or not a variety of staircases "looked climbable" fit very nicely with what biomechanical modeling determined to be the actual climbability of the staircases with respect to each subject. Subjects were at chance and had low confidence in their judgments when staircases were just at the limit of climbability, and subjects gave more and more accurate, confident judgments about staircases having both taller and shorter steps.
Warren determined that the maximum step height one can climb depends on one's ability to shift one's center of mass up over the foot with which one steps forward, which can be approximated by the step height at which one's leg would be maximally flexed upon stepping up. As it turns out, this means that regardless of the height of the subject, whether or not a staircase is climbable covaries with whether or not the steps are less than 88% of the subject's leg length and, as mentioned, Warren found a perceptual category boundary corresponding to this ratio (and, so, to the affordance of climbability).

Notice that this means the mentioned boundary is a proportion and so is intrinsically scaled. That is, it may be the case that your visual system need not come up with a representation of the independent height of the steps in order to track the climbability of the staircase. Warren (1984) suggests that we might, instead, track the climbability of the staircase by representing the steps as standing at some proportion of one's leg-length, as this information is available in the light as a proportion of one's eye-height (which is specified by things like the horizon, perspective information, or eye-in-socket rotation).

Though I've chosen stair-climbing as an example to discuss here, this line of research contains a great deal of excellent, similar work to which similar considerations apply. Mark and Vogele (1987) found judgments of maximally tall "sit-able" surfaces suggested a perceptual category boundary corresponding to the same biomechanically determined affordance base ("sit-able" surfaces can be no more than 88% of one's leg length). Warren and Whang (1987) found that (monocular, static) perceptual judgments of apertures as "pass-through-able" converge on a 1.3 ratio of aperture width to shoulder width. For more work toward identifying such body-scaled information, see Jiang and
Mark (1994) on gap-crossing, and for a nice summary of a variety of such work, see Warren (2007)

**Against the Fundamentality of Affordance Perception**

These results mean a number of exciting things. Warren's work and work like it seems to suggest that we perceive affordances, period, or at least suggest that it's not the case that our awareness of affordances is always the result of a very late, cognitive, conceptual judgment. The category boundary corresponds to a ratio the subjects were unaware of, and for which specifying light structure exists. However, support for the claim that the representation of affordances is *basic* is not supported by these results. Recall that affordances are *dispositional* properties of the environment. Certainly, we have reason to accept that when we look at staircases we enter into a state that (approximately) *covaries* with the climbability of the staircase. But we're also in a state that (approximately) covaries with a much more traditional property: the obtaining of a certain size-proportion between the height of the step and the height of the viewer's leg. Of course, this property is a relational one and, if we do use leg-height scaling to detect it, it involves viewpoint dependence, but the representation of this proportion is certainly not *thereby* the representation of the modal affordance property. One can, after all, believe all sorts of things about the lattice structure of salt without believing anything about what will happen if we put it in water (and vice versa). What reason do we have to think that we perceptually represent one property prior to processing to the other? It would seem Gibson still owes an argument for this, especially since, if what we're
perceptually representing is fundamentally the physical proportion and not the (assumed) covarying modal property, it's unclear how we might get to the representation of the modal property in any way other than by a sort of unconscious inference, which Gibson was at pains to avoid.

This suggests a general concern about the ability of results of this kind to support the claim that representation of affordance properties is basic. For any affordance, there must be some configuration of the world in virtue of which the affordance is there since, "affordances are constituted by material objects and their properties" (Warren 2005).

Whenever we have a perceptual state that covaries with an affordance, that state will also covary with at least some of the non-modal properties on which the dispositional property supervenes, and so far, we don't have any reason to think that our representation of the non-modal property is parasitic on our representation of the dispositional property and not the other way around.

Ultimately, which representations occur farther upstream in perceptual processing than others is an empirical matter. However, there is a methodological reason to favor the claim that the representation of traditional properties is basic. If we assume the representations of traditional properties is basic, we have a sketch of a story about the development of the capacity to represent the affordance properties. We have some idea of the way we take in the non-modal properties of the world (through constancy mechanisms, eye-height ratios, texture gradient information, and so forth), and of how we might learn laws linking some of these properties with the possibilities for action with which they approximately covary (say, through experience). That is, we have some idea of how representation of traditional properties might give rise to representation of
affordances by a kind of inference. But what sort of story can we tell if we assume
affordance properties to be representationally basic? It seems unclear what sort of story
might explain the representation of affordances without invoking prior representation of
traditional properties. It’s rather unclear how one might build representations of the
surfaces and spatial relations of the world out of modal representations like sit-on-ability
and pass-through-ability. This seems unattractive next to the alternative. Further, if
affordance perception is fundamental, the story about how we generate these basic
representations becomes complicated if it does not proceed upon representations of
properties like surface properties. There’s a reason ecological psychology typically
proceeds by looking for a configuration of the physical layout of the world that covaries
with the affordance in order to look for a perceptual state that covaries with that
configuration of the physical layout: affordances don't reflect light.

Because of the above, ecological psychology which proceeds by discovering these
covariance relations need not be taken to show that what we're really, at bottom,
representing perceptually is the affordance properties, but rather can serve as a line of
research which uncovers, on a case-by-case basis, a plausible physical base for
affordance properties and what the physical base does to structure the light so that the
information about it is visually available to us.

**Facing the Threat of Narcissism**

If the above is right, then the insights of ecological psychology can be turned in defense
of the “naturalist” against the concern that the narcissism of our sensory systems is reason
to think our sensory systems aren't in the business of producing states that are "about" anything. Akins makes a convincing case that sensory systems should be seen as the solution to an "engineering problem" (moving the organism toward mates and away from predators), but against Akins, the above considerations suggest that seeing sensory systems this way is compatible with seeing them as producing states that are "about" the objective world. Because any desirable or undesirable affordance property (that is of concern to our sensory systems due to narcissism, rather than a need to "report faithfully") depends on a configuration of physical, environmental properties, and because of the considerations above, it will always be available to us to say that what's being represented at the most fundamental level in such cases is that configuration of physical, environmental properties.

Of course, as mentioned earlier, Akins' concern about a move like this was that in claiming we represent affordance properties we've just gone from talking about a narcissistic system that doesn't represent anything to a representational system that represents only narcissistic properties, and that such a move won't help the "naturalist" strategy get off the ground. She writes:

"The naturalist’s hope was to find... a relation that would ground... a theory of aboutness. But in order to save the traditional view, the objector introduces narcissistic properties and thereby gives up on the link to the objective properties of the world. Prima facie, this is not a promising starting point for the naturalist theory.” (Akins 1996)

If the "narcissistic" properties are dispositional affordance properties, Akins is right to be concerned. As discussed above, if the representation of such properties is primary, the story about how we do such representing is rather mysterious. However, in light of the lessons I claim we should draw from the results of ecological psychology, Akins' worry
should seem less threatening. Ecological psychologists are, in each case (whether they would describe it this way or not), finding a relation between sensory states and external properties in a way that allows us to accept Akins’ contention that sensory systems be "for" guiding behavior while also dealing in "natural" properties. Ecological psychologists identify configurations of "natural" properties in the world which, if tracked visually, would allow us awareness of affordance properties at very little cost, and we have the makings for a good account of the way in which such configurations structure the light and put us into a perceptual state sensitive to that configuration. This program of research gives us reason to think it is in general possible to find some such configuration or set of configurations for any affordance property, even though doing so might be incredibly difficult in most cases.

Returning to Akins' particular thermoreception example, the thermoreceptive system is, indeed, not representing the temperature of the skin or the area surrounding the skin. It also does serve the function of keeping the organism within an acceptably safe range of temperatures and temperature changes. However, it is still available to say that it is doing so by putting us in contact with some configuration of properties of the world. The relational properties invoked in work like Warren’s (such as being 88% of one's leg height) aren't "narcissistic" in that they aren't defined in terms of "what is good for" the perceiver. They're merely objective properties of the environment expressed either in body-scaled units or ratios - on their own they have no dependence on properties based in the organism’s welfare or potential actions. When we look at an object intersected by the horizon line, we have present to us visual information about the height of the object in eye-height-units. The horizon line is always at one's eye level, and so the height of any
object, in eye-heights, is the same as the ratio of the vertical extent of the object in one's visual field to the vertical extent of the portion of the object below the horizon line. Nothing about this, however, is dependent on what might be good or bad for me, or on anything essentially private. If the horizon cuts a telephone pole in half in my field of view, the telephone pole is two my-eye-heights tall (about eleven feet).

Again, the claim is that our thermoreceptive system is putting us in contact with a configuration of properties in the world, even though it's difficult at present to say what that configuration is. It may, perhaps, be very complicated and may, perhaps, be disjunctive, but this is, if anything, a reason not to take it as an ideally simple case to generalize from, not a reason to think it must not be representing anything at all.

**Conscious Visual Experience of Affordances**

Finally, I move on to the last of the three earlier-mentioned issues which the priority of affordances might raise. This final section will discuss the content of conscious perceptual experience, and whether or not affordances are represented prior to, or at least independently from, "traditional" properties in conscious visual experience.

As mentioned, much recent philosophical attention has been focused on investigating what kinds of properties might be represented in conscious visual experience. Different theorists argue for a wide variety of candidate properties, from natural kind properties to the semantic properties of texts. It should come as no surprise that affordance properties are among those thought by some to be represented in conscious perceptual experience.
In Nanay (2012), Bence Nanay argues that our perceptual experience as of the affordance properties of an object is not parasitic upon our perceptual experience as of traditional properties like color and shape. Nanay makes a particularly compelling case, as his strategy is to avoid unreliable appeals to introspection as much as possible by arguing from empirical work involving unilateral neglect patients.

Unilateral neglect patients perform in an unusual manner with respect to a variety of vision-based tasks. They are slow or unable to find objects on their left sides if the objects are defined by a salient visual property ("point to the orange object", for example). When asked to produce a copy of a drawing, they will reproduce only the right half of the drawing. Nanay writes, "they are unaware of the shape, size, and color properties presented to them".

In Humphreys and Riddoch's (2001), the authors discuss the case of MP, who was a subject in a study in which the unilateral neglect patients were instructed to point at various objects described by features like color. MP did unusually well in some of the trials and, when questioned, explained that he sometimes was able to find the targets if he "thought about what to do with them" (ibid.). So, in subsequent trials, MP and two other subjects were asked to point out objects by color, by kind ("point at the cup"), by functional description ("point out the thing you drink with"), and in response to a gesture made by an experimenter (a hammering motion to indicate the subject should point to the hammer, for example). MP did better when directed by gesture or by functional description and, interestingly, MP did worse again when directed by functional description when the object to be pointed to was oriented such that it could not be used as described (MP had a lower
success rate in pointing to an upside-down cup when prompted to "point to the thing you drink from", for example).

Nanay argues that conscious experience of affordances cannot be grounded in experience of traditional properties, as MP is a disproving counterexample. He writes,

"the empirical evidence from unilateral neglect has a very straightforward consequence: it is possible to have an experience of action-properties without the experience of shape and color properties - that is exactly what happens during the visual search tasks of these unilateral neglect patients. Hence, the experience of action-properties cannot be based on, or grounded in the experience of shape and color properties." (Nanay 2012)

There are a number of places, however, at which one may resist this conclusion.

First, there seem to be a number of considerations pushing against characterizing unilateral neglect patients as having no visual experience of things in their left field of view. A prominent opinion in the literature based on the neuroscientific evidence seems to be that the impairment suffered by at least some cases of unilateral neglect is an impairment in "exogenous" or stimulus-driven attention on the left side (Bartolomeo and Chokron 2002). These patients respond to visual stimuli on the left side significantly better when their attention is explicitly drawn there, especially when there isn't anything going on in the right side to distract them. In fact, in the very Humphreys experiment Nanay cites, there was a condition in which all three unilateral neglect patients were asked to name the object and its color in a single-object display, which all three subjects had no difficulty in doing. It was only when given a multiple-object search task that they began to perform poorly.

---

4 I have recently found out that Farid Masrour made a point similar to the one in this paragraph (citing different support similar in kind) in a reply to a talk Nanay gave on his (2012).
Humphreys et al seem sensitive to this, speaking mostly in terms of investigating what kinds of visual "search templates" we might be equipped with. Humphreys and Riddoch claim in their (2001) merely that "action templates can influence visual search and selection independently of perceptual templates of targets" and they discuss MP's good performance on action-cued searches in terms of "facilitation of the process of matching a stimulus to a search template", both of which seem quite compatible with enjoying the ordinary swath of color and shape visual phenomenology but having an impaired ability to match such content to some form of stored representation.

Further, other work by Humphreys and Rumiati actually provides some positive reason to think that experience of functional or affordance properties is grounded in experience of color and shape properties. In Humphreys (2000), Humphreys discusses an experiment in which ordinary subjects are shown pictures of objects and required to respond very quickly (before a tone is played) with a gesture appropriate to the displayed object. A relatively high proportion of errors were errors that were related to the shape properties of the pictured object - making a hammer-swinging gesture in response to a picture of a razor, for example. Assuming, however, that the idea is that the short time limit is preventing later iterations of processing (that is, if the short time limit helps us get at the "ground level" of visual experience), if the visual experience of functional properties were not grounded in the experience of shape properties, we might expect to see function-based errors, not shape-based errors. Presumably the participants made a hammer-swinging gesture in response to the razor because the razor silhouette has a similar shape to the hammer silhouette, but if visual experience of functional properties were not grounded in experience of shape properties, you might expect, say, other carpentry related errors, or
perhaps motions one might make with other hammer-like tools, and so on. Granting some assumptions about hierarchical processing of visual experience and time constraints, the best explanation of the high proportion of shape-based error seems to be that there is a "bottom level" of visually experienced shape and color properties upon which visual experience of affordance properties, if there is such a thing, is based.

Nanay, however, cites a personal communication with Glyn Humphreys in support of his claim that unilateral neglect patients really do lack color and shape phenomenology on their left sides. In addition, unilateral neglect symptoms can vary from case to case, so let us for the moment grant that unilateral neglect patients really lack color and shape experience on their left sides. It still seems that one might continue to resist Nanay’s interpretation by denying that subjects like MP enjoy a visual experience as of affordance properties at all.

Nanay anticipates this sort of response, writing,

"What about their perceptual phenomenology then? The objector is forced to conclude that these patients lack any perceptual phenomenology while they are performing this visual search task. The only properties they are aware of are action-properties, but these properties are, by supposition, not part of their perceptual phenomenology. This is an extremely problematic conclusion as these people are staring at objects, perform visual tasks with what they see, talk about what they see, manipulate what they see, and, importantly, consciously experience what they see, nonetheless, the objector needs to say that they lack perceptual phenomenology: there is nothing it is like for them to see these objects." (Nanay 2012)

I will admit that I'm not quite sure what Nanay means by asserting in this list that they "consciously experience what they see". I would have thought the other items in the list (like that they manipulate what they see and talk about what they see) were bits of evidence that they consciously experience what they see. But if I'm not misunderstanding, the complaint is that it is implausible that the unilateral neglect patients suffer a
phenomenological void. In response, we might compare the situation of the unilateral neglect patient to that of the blindsighter.

If unilateral neglect patients lack experience of traditional properties, this would put them in a rather similar position to that of blindsighters, who lack conscious awareness of things in their visual field despite displaying a number of curious proficiencies such as being able to correctly grasp objects in front of them when prompted, and being able to make correct "guesses" about the orientation of a line or a grating. Milner and Goodale (1995) explain blindsight in terms of their famous ventral/dorsal system distinction. Carruthers (2001) reviews their findings, writing that a good amount of visual information can be available in the parietal cortex (thought to be involved in the control of movement) even when none is available in cortical area V1 (thought to be connected with conscious experience). The result is that patients with damage to V1 may have a great deal of visual information available for the control of movement, but of which they have no conscious awareness. The reason, suggest Milner and Goodale, that blindsighters are capable not only of grasping and pointing, but also of talking about the things in their field of view is that the subjects may be priming themselves with motor-programming cues, like tracing the orientation of a grating with the hand or eyes. (Carruthers 2001)

I suggest that the same may be the case for our unilateral neglect patients. Nanay warns against invoking Milner and Goodale (1995), writing that if the visual information about affordances is relegated to the dorsal stream, "[this scenario is not] consistent with the suggestion that the dorsal stream can give us conscious experience of what an object can be used for in the complete absence of any conscious experience of its shape and color" (Nanay 2012). However, the claim here is not that the content of the dorsal stream
representations give rise to a visual phenomenology. Rather, the claim is that dorsal stream representations of traditional properties may give rise to motor representations that are access conscious but not phenomenally conscious, which might then explain the behavior driving Nanay’s view (interacting with objects, talking about them, and so on).

In summary, it is doubtful that perceptual representation of affordances is prior to representation of traditional properties. Empirical results showing only that we track the affordance properties cannot support such a claim, and there are methodological reasons to lean away from it. I have also argued that unilateral neglect cases do not show that experience as of affordance properties is independent of experience of traditional properties, since, in the first place, there is significant reason to think that unilateral neglect patients do have low-level visual phenomenology and, in the second place, even if they do not have low-level visual phenomenology, we can explain away what Nanay takes to be evidence that the unilateral neglect patients do have visual experience as of high-level affordance properties in the first place.

Affordances in conscious perceptual phenomenology is but one of many controversial topic in the general discussion of rich perceptual phenomenology. In the next chapter, I turn to address several well-developed others, as well as arguments for rich content generally, applying some of the lessons drawn from the case of affordances here.
Chapter II: Rich Content in Visual Experience

For most of us, daily life is a whirlwind of conscious perceptual experience: you might currently, for example, be enjoying the smell of your favorite coffee house or perhaps the sound of your best friend's voice or the look of the colors in your brother-in-law's latest efforts at graphic design. Ordinarily, in cases like these, each of these things has a characteristic perceptual "feel" - a perceptual "character" sometimes designated with the phrase "what it's like". There's said to be "something it's like" to sniff fresh coffee. "What it's like" to smell fresh coffee, is very unlike "what it's like" to sniff fresh fish, and it's at least intuitive to think it would be quite different in this way no matter what your beliefs or judgments were about what you were smelling. The character of this conscious perceptual experience or, as I will call it from here on out, "perceptual phenomenology", will be the target of this section. Specifically, this chapter will consider what kinds of properties might figure into the content of our perceptual experiences, with a special focus on visual phenomenology.

To illustrate what I mean by "content", consider the visual experience you are having right this moment as you read this chapter. There is an intuitive sense in which your visual experience might be thought to be accurate or veridical only if there are, in fact, white rectangles in front of you with a particular arrangement of letter-shaped black blots on them. Crucially, in addition to this, it's very intuitive to think your present visual
experience has the accuracy conditions it does, in some sense, *in virtue of* the "what-it's-like" character of your present visual experience: if your visual experience had had a different character, the accuracy conditions would have differed as well in some sort of appropriate way, and vice versa (it is possible to deny this, but I will assume it in what follows⁵). In this sense, we might say, your current visual experience has certain *representational content*. Your current conscious visual content is such that it represents that there are white rectangles with certain black blobs on them in front of you and is veridical only if this is the case, and that this content is at least very closely tied to visual phenomenology. So, given this framework, in inquiring into the kinds of things that are represented in visual experience, we are inquiring into visual phenomenology itself as well.

It might seem, however, puzzling at this point why we should have any difficulty at all determining what kinds of things figure into our visual experiences. After all, in considering your present visual experience in reading this chapter, all you had to do was, perhaps, introspect a bit, and the phenomenological character just seemed evident: it just *looks* like there are white rectangles with black blobs near me, so my visual experience must represent there being white rectangles in with black blobs near me.

Perhaps in the extreme cases, it might be hard to disagree about what kinds of properties introspection reveals as figuring into our perceptual content. On one hand, when I look at my new neighbor's mailbox and notice that it's stuffed full of Oxfam mailers, although I might say "I can see that the new neighbor is generous," it seems quite

---

⁵ One can, of course, hold that visual phenomenology and representational content come apart, or even that our visual phenomenology has no content at all. This is an interesting area already rich with philosophical discussion, but since my purpose here is evaluating "Rich Content" style views which assume a close tie between visual phenomenology and representational content, I will do the same in what follows.
clear that my neighbor's generosity is not really represented in my visual experience (certainly not when I'm only looking at his mailbox). Rather, it seems clear that my visual experience involves something having to do with the shapes and colors of the mailbox and its contents and it is on that basis that I make a conscious cognitive judgment about one of my neighbor's properties. Similarly, on the other hand, it seems hard to deny that color is a part of the content of ordinary visual phenomenology - it seems pretty clear on introspection that I don’t have a raw non-colorful visual experience on the basis of which I'm simply making a cognitive inference about colorfulness. Color just seems to be "right there" in my visual phenomenology.

However, there are more difficult cases that introspection fails to settle. Consider the following image:
Introspect into your visual experience as you look over the image, and try to take note of its character. Now consider the following: the image is an image of a horse and rider in profile (the head is at the top left, and the horse has three hooves on the ground and one aloft). Now, reexamine the image. Most viewers (myself included) find themselves strongly pulled toward the claim that they enjoy a very different visual phenomenology after having learned about the horse and rider. There was a change in visual experience - the image looked different somehow. But what could such a difference in visual experience amount to? I must have had an experience as of just the same shapes and colors in the same positions after learning of the horse and rider as I did beforehand. After all, I just looked at the very same image twice. Still, there seemed to be some kind of visual difference between those two experiences.

This kind of case has led to the increasingly popular suggestion that our visual experience sometimes represents more than just shape, color, motion, and other such low-level properties standardly taken to figure into perceptual content. I will follow Susanna Siegel in calling this the "Rich Content View":

**Rich Content View**: In some visual experiences, some properties other than spatial properties, color, shape, motion, and illumination are represented. (Siegel 2010)

To see its appeal, consider how, in the case above for example, such a view might seem to have an explanatory advantage. It seems quite intuitive to say that when you looked at the image the second time, the reason it looked different is that a new property was represented in your visual experience: the display appeared as a horse and rider, whereas the first time it did not - it was merely an arrangement of dark blobs. Siegel uses the following toy example to more clearly differentiate the kind of claims rich theorists
make: Imagine that you visit a friend's home and come upon a bowl containing piece of expertly-made fake fruit, completely visually indistinguishable from ordinary fruit (ibid.). A rich content theorist would not claim that you have an experience as of some shapes and colors and then proceed to make an incorrect high-level inference to the effect that there is ordinary fruit on the table. Rather, the rich content theorist might claim that you visually experience what's in front of you as fruit; the fruit-ness on the scene figured into your conscious visual experience, and you just came to believe things are exactly the way they looked. That what's in front of you is, in fact, fruit (and not wax) was part of the character of your visual experience just like the blue-ness of a clear sky or the smell of baking bread. As such, in the fake fruit example, the rich content theorist would claim that you're having a non-veridical perceptual experience as of real fruit in a bowl - that is, your experience is illusory (whereas a non-rich theorist, by contrast, would say the experience is perfectly veridical, it's just that you formed a false belief on its basis).

Is it really evident on introspection which way things are with the gestalt horse and rider, for example? We've noted reasons that might tempt one to adopt a rich-content explanation, but might it not also be the case that I have a certain visual experience as of black blobs and enter into a high level cognitive state on the basis of that experience instead of there being, so to speak, a horse-and-rider-y character to my conscious visual experience in addition to the belief I have about what's depicted in the image? Perhaps you have strong feelings one way or the other, but in the literature the results of introspection have been mixed. Different discussants insist the results of introspection return different results in a number of cases. In light of this, our investigation into the
kinds of properties figuring into the content of visual experience will focus on avoiding sole reliance on introspection to settle the matter.

First, a few clarificatory notes are in order. For one, rich content theories need not be limited to any particular set of high-level properties – “rich” properties are merely those that traditional accounts of visual phenomenology have tended not to include. There is nothing, for example, riding on any unifying "richness" that might sort perceptual theories into those that are truly "Rich Content" and those that aren't. The label merely collects theories that include an endorsement of the rich content view, as stated above. The rich theorists may discuss any of a wide variety of non-traditional properties in arguing for the “rich” view. Suggestions, for example, have included kind properties (as with the horse and rider above), aesthetic properties (the beauty of the painting seems to make a difference to perceptual phenomenology), semantic properties (text seems to look different before and after one learns the language it's written in), moral properties (John has looked different to me ever since I discovered he was a murderer), and functional or affordance properties (doorways I fit through have a different "look" from doorways I don't fit through, above and beyond their different shape) several of which were discussed in the previous chapter.

Another note: it is rather unclear how to arrange properties into a suitable hierarchy of "richness", but for reasons similar to those mentioned in the previous paragraph, we can safely put that matter aside here. Rather than worry about "how rich" visual experience gets, the important thing to consider here, as is evident in Siegel's formulation of the view above, is simply whether or not any properties other than spatial properties, color, shape, motion, and illumination are represented in visual experience.
These properties I will call "traditional" or "low-level" properties, and all other properties than these I will call "rich" or "high-level".

If it turns out that any rich properties are represented in visual experience, that alone would be quite philosophically exciting. There are a number of reasons for this. For one, it is simply intrinsically valuable to converge on the right account of the content of visual phenomenology, hopefully explaining the disagreement in the process. To be able to agree on a principled account of the features of visual experience is part of the primary business of philosophy of perception. In addition, however, the properties that are represented in visual experience might turn out to make a massive epistemological difference. If it turns out that perceptual experience has any sort of epistemically privileged position with respect to the conferring of justification, for example, then what properties are represented in experience might drastically change what we're justified in believing. This might also, in turn, make an important moral difference - depending on your moral theorizing, a person that was having a hallucinatory visual experience as of a man aiming a gun at her might turn out less blameworthy for behaving violently than a person who had a veridical visual experience as of a man with a black blob in his hand which she subsequently inferred to be a gun.

Further, as Siegel points out, we'd also like to be able to have a good account of representation itself. In particular, we'd like to know in virtue of what do our mental states come to represent states of affairs in the world - what relation is it that explains the way in which my beliefs about my spouse are about one person and not another. A very common approach to this kind of investigation is to start by investigating perceptual content, since this is a relatively well-understood and hopefully a relatively simple case
of representation\textsuperscript{6}. But if we have independent reason to suspect some properties and not others are represented in visual experience, this will constrain the kind of theories of representation we can accept. As Siegel puts it, "[i]f visual experience can represent that there are peaches on the table, then whatever makes it the case that a visual experience has the content it does had better not disallow that visual experiences represent the properties of being a peach or a table." (ibid., p.8)\textsuperscript{7}

In what follows, I will argue that there is less reason to think the rich content view is true than it might seem in light of the view's flagship cases. In these cases, there is a perfectly sufficient low-level alternative. I will first discuss ways of investigating the matter and discuss reasons to move away from relying on introspection to settle things. I use Siegel's "phenomenal contrast" method, a method intended to move discussion away from reliance on introspection, as a launching point for this discussion. I then argue that evidence for each kind of "rich" property must be assessed on a case-by-case basis, and then proceed to evaluate four prominent cases made for the rich content view. These are, in the order I will be considering them, an argument for the inclusion of kind properties from attention/categorization research, an argument for the inclusion of face properties, an argument for the inclusion of the property of animacy, and an argument for the inclusion of causation-related properties. In each of these cases, I argue that "high-level" theorists fail to rule out a "low-level" explanation of the experimental results discussed, and since, as I will argue, this is the task the high-level theorist needs the experimental

\textsuperscript{6} For discussion of another intersection between accounts of perceptual content and accounts of representation generally, see “Sensory Narcissism” and “Facing the Threat of Narcissism” in Chapter I.

\textsuperscript{7} For additional ways in which the Rich Content View, as framed above, is relevant to other inquiries, see Siegel 2010's introduction under the section "Why Does it Matter?"
literature to accomplish, we are left with little reason, at present, to think the rich content view is true.

**Introspection and the Phenomenal Contrast Method**

Siegel nicely lays out the case against relying solely on introspection to settle what properties figure into visual experience, writing:

"[E]ven if we grant that phenomenal character is intrinsically accessible to introspection, it does not follow that introspection tells us sufficiently precisely which contents such states have." (ibid.)

Given the sharp disagreement that exists in the literature (together with the fact that it's rather unlikely that theorists are having different visual experiences: some with high level content and some without), this seems quite plausible, and seems to be a generally commonly held view (Prinz 2006, for example, makes a similar point in discussing introspectionist psychology, see also Nanay 2012, Block 2014). Siegel writes that if introspection could tell us who was correct, we should find no disagreement among theorists and rightly notes that it's even rather unclear what one ought to do when told to introspect.

Somewhat more substantively, she also points out that introspection is plausibly domain-general, allowing one to "discern that one is currently in a mental state, without discerning exactly what kind of state it is" (Siegel 2010, p.82). She writes:

"[E]ven while introspecting correctly, one could unknowingly introspectively attend to something other than visual experience. Given this possibility, verdicts delivered by introspection that purport to be verdicts on the contents of visual experience may in fact be about the contents of some other kind of state, such as a judgment."
Because of this, Siegel urges that we instead use her method of *phenomenal contrast* in order to ease off of the reliance on fine-grained introspection in trying to discover whether or not the rich content view is true. This method involves considering two overall conscious experiences (total experiential states, including anything the theorist might count as making any phenomenal difference, including visual experience, auditory experience, cognitive phenomenology, background conscious states, etc.) which differ phenomenally in *some* way, and inquiring into the best explanation for that phenomenal difference.

The advertised advantage to this method is that, rather than using introspection at a very fine level to try to discover *what* the contents of one's visual experience are (which Siegel has argued is a task for which introspection is unfit), the phenomenal contrast method only requires that introspection deliver the shared intuition that the two proposed overall experiences would differ phenomenally *at all*. Siegel writes:

"[T]he kind of intuition on which the contrast method rests is simple and modest. Such intuitions concern whether there is a change in phenomenology between two sorts of situations. It would be quite radical to deny that there were any such cases in which introspection could detect a phenomenal contrast between overall experiences." (ibid., p.91)

The process then aims at settling things by inference to the best explanation. Making the case *for* the rich content view would be to describe two overall experiences which all can agree differ in *total* phenomenology, and then to argue that the best explanation for that difference is that in one case the subject's *visual* experience represents a high-level property and in the other case not. A counter-argument will typically take the form of a rival explanation: the thin theorist admits that there *is* a difference between the two total experiences, but argues that that difference is best explained by a difference in, for example, non-perceptual phenomenology. The rich content theorist will then typically
modify the cases in an effort to rule out the proposed rival while leaving the rich-content-involving explanation intact. In this way, investigation can proceed on the basis of more modest, general introspection. This sort of work is the majority of this section of Siegel's book.

Consider the following as an example of this process: In arguing that kind-properties figure into the content of visual experience (in this case, the property of being a pine tree), Siegel proposes the following pair of scenarios:

"Suppose you have never seen a pine tree before and are hired to cut down all the pine trees in a grove containing trees of many different sorts. Someone points out to you which trees are pine trees. Some weeks pass, and your disposition to distinguish the pine trees from the others improves. Eventually, you can spot the pine trees immediately: they become visually salient to you. Like the recognitional disposition you gain, the salience of the trees emerges gradually. Gaining this recognitional disposition is reflected in a phenomenological difference between the visual experiences had before and those had after the recognitional disposition was fully developed." (ibid.)

Siegel has proposed two scenarios, one in which you're seeing pine trees for the first time and one in which you've acquired a recognitional disposition through experience, and requires only that all agree that the person in the two scenarios would, indeed, have some difference in total phenomenology (visual or not, perceptual or not, etc.). She then goes on to claim that the best explanation for the difference is that there is a difference in visual phenomenology: in the early case, you do not have a visual experience as of pine trees (perhaps only green and brown shapes and other traditional properties) and in the later case, you do have such a visual experience.

An opponent of the rich content view should now argue that an alternate, better explanation exists. It seems to me, for example, that a better explanation for the phenomenological difference between the cases is that the experience with pine trees has
changed the way your visual search and recognition templates work, and in what groupings you allocate attention to the low-level features in your visual experience, and these do result in changes to "what it's like" upon approaching the trees as an expert but have nothing to do with visual phenomenology which includes pine-tree-ness. That is, even though this explanation admits of a change in visual phenomenology, this is merely a change in low-level features, and so to hold that we also represent the pine-tree-ness in the scene in front of us would be unnecessary. The proponent of the rich content view might then respond by arguing that my proposed alternate explanation leaves something out and attempting to show this by producing a new pair of cases which also differ in total phenomenology, but which exclude the thin theorist’s proposed explanation. This process, the rich theorist hopes, makes increasingly apparent the kind of phenomenological difference that the proposed alternative supposedly fails to account for.

This way of investigating the content of visual experience has produced a number of interesting points on both sides, and there has already been quite a bit of back and forth. For readers interested in the particulars of this back and forth (for example, see Prinz 2013, Campbell 2013, Siegel 2013c). However, rather than defend my preferred explanation for particular pairs and add it to the pile, I will here be more concerned with arguing that there is reason to think such a process, by itself, is in general still just as open to over-reliance on introspection as the usual methods of analysis. Ultimately, I argue, what is important is that the rich theorist attempt to find cases where the presence of high-level content is independently motivated by empirical work.

Certainly, it seems true that introspection can at least ordinarily reliably tell us whether or not there is a difference at all between two total experiences, but the phenomenal
contrast method will actually require more than that to get off the ground in any interesting cases. The more specific the cases get, the less of a trivial feat of introspection this will be. Consider the way things will typically go: the first theorist proposes two hypothetical total experiences, hoping that the best explanation of the intuition that they differ phenomenologically at all is that they differ in the theorist's preferred way (in this case, the theorist hopes the best explanation is a difference in rich content). Those opposed will propose alternate, low-level explanations. But insofar as the alternate explanation is at least a reasonable account of some non-rich-property phenomenological difference between the two scenarios (and these differences are always plentiful), the alternate explanation will be sufficient for accounting for the mere fact that the two total experiences are agreed to be different in some way. The problem is that the supposed common ground between rich and thin theorists, the fact that the total experiences differ at all, will almost certainly be massively overdetermined: there will be many ways in which experiences in the two scenarios might reasonably differ. The rich content theorist needs to argue not that the alternate explanation is a worse explanation, but rather that the thin theorist has the wrong explanandum. However, to be able to claim something like this, the explanandum can't just be the fact that there's any difference at all. One would have to have be able to single out some particular difference in content between the two total experiences and be able to determine that the alternate proposal fails to explain this difference. But the fact that we had no reliable way of introspectively determining the specifics about differences in content between two total experiences was the problem the phenomenal contrast method was supposed to assist us in overcoming in the first place. The more highly specific the rich theorist’s proposed total experience pairs are, the more introspection and intuition they
To put it another way, in order for the rich theorist to argue the thin theorist’s explanation is a worse explanation, she must generate a pair of contrast cases between which there is strong \textit{independent} reason to think a difference in represented "high-level" properties exists. Certainly, the phenomenal contrast method suggests the structure of an argument the rich content theorist might wish to try to run, and is unquestionably of value in framing the debate by highlighting the key element that rich theorists and thin theorists ought to discuss in order to determine whether or not the rich content view is true. The phenomenal contrast method makes it clear that we ought to go in search of pairs of cases there is independent reason to think must differ with respect to high level content. However, the phenomenal contrast method will not be of help in \textit{finding} such cases, and will not help us determine, of any particular pair of cases, whether or not we have that \textit{independent} reason to think there is such a difference between them.

Since the phenomenal contrast method itself only helps us frame the problem at hand, what should we do to search for such cases? Although intuitions about cases of purported high level perceptual content are vulnerable to alternate, low-level explanations, as we've seen, we needn't stop there: we ought to try to make systematic, controlled attempts to rule out these low-level explanations on a case-by-case basis, using the experimental literature when relevant, as success in this endeavor would put just the kind of \textit{independent} support behind the rich content view mentioned above.

So, let us proceed by examining some of the best cases made in, and based on, relevant empirical work several different high-level properties.
Case 1: Inattentive Categorization

Let us begin by considering William Fish's suggestion that certain kinds of evidence from attention research might militate in favor of the rich content view (Fish 2012).

Fish describes research conducted by Li et al (2002) in which participants were instructed to fixate on a cross in the center of a display which was replaced with an array of 5 randomly rotated letters at the onset of the experiment. The 5 letters were either all the same, or else 4 were the same and 1 was different. Further, 53 ms into the experiment, a peripheral stimulus (an image drawn from a varied pool) appeared at the corner of the monitor for 24 ms. One group of subjects was asked to ignore the peripheral stimulus and determine whether or not all five letters appearing in the center were identical and indicate this with a button press. Another group was asked to ignore the central stimulus (the 5 letters at the fixation point) and quickly determine whether or not a peripheral stimulus fell into a particular category. In the third condition, participants were asked to perform both tasks simultaneously. The idea in this third condition was that when participants were required to perform both tasks, the center-task would consume the participants' attention and so their performance on the periphery-task would give us information about our ability to perform visual sorting tasks preattentively.

When the subjects engaged in the periphery-only task, they were about equally accurate (around 75%) across all the categories in the experiment: they were equally proficient at determining whether or not the stimulus was a certain letter, whether or not the stimulus was a disc of a particular arrangement of colors, and whether or not the stimulus contained images representing various recognizable kinds (such as animals and
cars). However, surprisingly, when subjects performed both tasks simultaneously, performance on the periphery-task fell to chance when the category was color or letter-based, but remained the same when instructed to determine whether or not the image contained objects of a certain kind.

Fish argues that this should count as at least some reason to think the rich content view is true. Fish is somewhat reserved here, and the section is short, but the following is, I take it, the best reconstruction of the argument:

1. In the Li et al experiments, subjects are able to successfully categorize peripheral stimuli as animals while simultaneously engaged in a perceptual task at the center of the visual field.

2. Thus, the Li et al subjects are able to successfully categorize peripheral stimuli as animals without attending to them.

3. Post-perceptual processing requires additional resource allocation and, therefore, requires attention.

(from 2 and 3) The categorizing of the peripheral stimuli as animals does not occur post-perceptually.

Thus, the categorizing of the peripheral stimuli occurs perceptually.

If we accept this, the mentioned experiment would seem to suggest that some kind properties do figure into visual experience, whereas some pattern and shape-recognition properties do not. After all, it was difficult to sort on the basis of shape and pattern properties without attention, but participants seemed quite capable of sorting on the basis of kind properties without attention. This would be a surprising and interesting result, to say the least!

There are, however, reasons to be suspicious of this argument. For one, the move from (1) to (2) is questionable. The claim that the task at center made it such that no
attentional resources whatsoever were dedicated to the stimuli at periphery is a rather strong claim - one we would need reason to think is true, since it is far from obvious. Note that even Li et al seem not to be willing to assert that the experimental paradigm they used reliably results in completely unattended peripheral stimuli, writing that they've shown that the categorization occurs in the presence of "very little or no focal attention" (Li et al, 2002), leaving room for the possibility that the central task either did not completely consume participants' focal attention or that peripheral processing may have benefitted from participants' non-focal attention.

Further, (3) at least stands in need of support. Fish holds that it is "independently plausible [that] for processes of interpretation to take place, we would need to allocate additional cognitive resources to the task". However, even if we accept this plausible claim, it doesn't entail that the additional cognitive resources must take the form of focal visual attention. Fish’s argument can only proceed if the kind of attention required for the central visual task excludes the possibility of providing the kind of attention necessary for categorization. This is more demanding than the plausible claim that categorization requires attention at all, and in Fish's admittedly short and merely suggestive discussion, this claim isn’t provided direct support.

**Case 2: Animacy**

A second prominent case for rich perceptual content gaining recent support is the case for “animacy perception”. This is, roughly, the case for having perceptual representations as of agents, agency and, perhaps, as of agents' intentions as well. Being able to identify other
agents and being able to predict their behavior quickly and reliably is of clear importance to the survival of creatures like ourselves. It's easy to see how adaptive pressure might favor a dedicated, low-level agent or agent-related-property perceptual system⁸. As a result, there is a great deal of empirical interest in the area.

The animacy empirical literature is rich, with no shortage of investigation into visual content conducted on the basis of introspection and report. The results are certainly suggestive (in fact, Scholl and Gao 2013 suggest this introspection report data is likely the driving force behind the entire program of research). Consider, for example, the classic Heider and Simmel animation consisting of several simple moving polygons and lines. Viewers asked to freely describe what they've seen seem to find describing the animation in terms of mental states and intentions irresistible (e.g. "the large triangle is angry with the small triangle, and the small triangle is trying to get away"), and in fact have quite a bit of difficulty describing the scene when instructed not to use mental-state terms (ibid.).

Gao, McCarthy, and Scholl (2010) produced an ingenious refinement to demonstrations like the Heider and Simmel animation which, interestingly, makes quite a bit of progress toward isolating some of the properties of a visual display that might be sufficient for producing the kind of result mentioned in the Heider and Simmel case above. Gao et al produced an interactive demonstration in which a user-controlled disc appears on a screen littered with pointed dart-shapes moving about in a fixed pattern (unrelated to the user-controlled disc). In one condition, the darts constantly rotate to face the user-controlled disc, even though their translational movement is unaffected by this rotation. In free report, participants described the display in ways invoking intentional activities or

---

⁸ Compare, for example, J.J. Gibson's reasons for thinking we ought to begin by looking for evidence of things like threat-perception, hiding-place-perception (see Chapter I of the present work).
mental states similar to the Heider and Simmel case above (e.g. "There are many arrows chasing after my circle"). However, in another condition in which the arrows had identical translational movement, but constantly rotated to stay pointed 90 degrees away from the user-controlled disc, participants instead described the display without invoking such terms (e.g. "there were arrows that went in random directions" or "chaotically floating white chevrons"). Scholl and Gao 2013 take this to be evidence of a genuine perceptual effect, or as they put it:

"[The display with darts facing the disc] evoked a rich percept of animacy and intentionality and [The display with darts facing away from the disc] did not. This was clear from the resulting phenomenology (as readers can experience for themselves via the demos at http://www.yale.edu/perception/wolfpack/), but it was also clear in terms of free reports."

Additionally, Scholl and Gao mention that similar experiments using diverse populations of participants seem to get the same result across the board, excepting only participants such as those with disorders of social perception, such as autism spectrum disorder or damage to the amygdala. Scholl and Gao (2013) don't expand much on in what way, exactly, they think this univocal result across highly diverse groups supports the claim that animacy figures into our perceptual content, but a reasonable extrapolation would be that the univocal result they mention militates against the reports' being the result of a high level, conceptual judgment, as one would expect very diverse populations to have very different conceptual resources and so we ought to, accordingly, expect different reporting behavior if the reports were the end result of a such judgments\(^9\).

\(^9\)Though I do not wish to attribute this line of reasoning to Scholl and Gao, as they may have been intentionally conservative here. I mean only that this is a reasonable line of thought proceeding from the findings they mention which it will be valuable to consider.
However, we have already seen reasons that we ought not rely on introspection and introspection-based reporting to investigate the content of visual experience. Scholl and Gao hold a similar position. They write:

“Of course, phenomenology alone cannot settle the question at hand... Indeed, perhaps the central lesson of cognitive science is that our introspective intuitions about how the mind works are fantastically poor guides to how the mind actually works.” (Scholl and Gao, 2013)

Thus, as they point out, they mention the above results only to suggest that animacy is a good candidate for a genuinely perceptual high-level property which ought to be given a more principled investigation and analysis which, fortunately, they claim to provide. This stronger argument in favor of animacy perception proceeds along three main lines of argument proceeding from their work in empirical psychology. I will argue that these three lines do not amount to a strong argument that representing animacy is perceptual at all, and so, accordingly, do not make a strong argument that animacy properties figure into visual phenomenology (though, of course, Scholl and Gao’s work is of deep interest for a plethora of other reasons).

First, Scholl and Gao mention a rather amazing group of studies they put under the heading “chasing” studies. In particular, participants in the “don’t-get-caught” task were in control of a green disc (the “sheep”) on a monitor surrounded by randomly moving white discs. Hidden amongst the randomly moving white discs was a special white disc (the “wolf”) which, if it came into contact with the “sheep”, would end the task in failure. The task ended in success only if the participant could manage to maneuver the “sheep” away from the “wolf” for a set amount of time.

As it turns out, participant success could be quite reliably controlled by various adjustments to the “wolf”’s behavior. In one condition, the “wolf” was constantly moving,
but in each trial the experimenters varied the amount of angular variation away from a beeline for the “sheep” which the “wolf” was allowed. The experimenters called this “chasing subtlety”. For example, a wolf with a chasing subtlety of 0 degrees was allowed to move in no direction other than directly toward the “sheep”, whereas a “wolf” with a chasing subtlety of 90 degrees was allowed to move in any direction up to a right angle from the direction of the “sheep”, but never away from the “sheep” at all. It turns out performance on this task is a U-shaped function with the best performance rates occurring at very high and very low chasing subtlety, and the worst performance occurring at middling chasing subtlety. This was because actual chasing and chasing detection come apart, and a successful wolf chases, but is not detected. At the extremes, the “obvious” wolf that makes a beeline for the sheep is easy to avoid because he’s easily identifiable as the threatening disc, whereas the “incompetent” wolf that moves entirely randomly is impossible to identify as the threat, but fails to do any actual chasing of the sheep. (Gao, Newman, and Scholl 2009)

Another condition mentioned had participants attempting to avoid a wolf whose beeline chasing behavior was punctuated with random movement at various temporal intervals. Gao and Scholl were able to obtain similar results – success can be reliably controlled by varying the percentage of the time the wolf spends making a beeline for the sheep, and Gao and Scholl found a similar U-shaped success function with highest relative successes at 0% “obvious” movement and 100% “incompetent” random movement (Gao and Scholl 2011).

These results, say Gao and Scholl, count in favor of genuine animacy perception, writing:
“The point of reviewing these studies here is to note that chasing detection (as a form of perceived animacy) is influenced in systematic ways by rather subtle display parameters, in the form of a psychophysical function… We may have some categorical intuitions about the factors that should count as animate, but we surely do not have intuitions about the difference between 30 degrees and 60 degrees of chasing subtlety… (Indeed, it is not always even possible to tell the difference between such displays during free viewing, and it is difficult to make a conscious decision about animacy on the basis of information that you are not even aware of.)” (Gao and Scholl 2013)

Participants, they reason, could not, for example, merely be behaving in accordance with what they believed to be reliable signs of animacy, since it’s implausible that any participants brought to the experiment antecedent beliefs about what percentages of time spent in random movement did and did not reliably indicate chasing (that we have no such beliefs is why their results are so fascinating). Further, they reason, even putting this aside, the behavior of the participants couldn’t be given a high-level cognitive explanation since participants were surely not even aware of the subtle differences in chasing behavior that controlled their performance.

The main problem with this line of argument, however, is that it seems perfectly plausible that there might be a variety of behaviors which vary in accordance with very subtle visual properties, and about which we have no conscious beliefs or expectations, but which are nonetheless mediated by high-level conceptual processing. Consider, for example, that you in general visually recognize and behave appropriately toward things that are cars and things that are not cars, even though you likely can’t say precisely what visual factors make the difference between your taking something to be a car and not. It seems perfectly plausible that your recognition of cars is guided by these subtle visual details despite the fact that it seems fairly clear that a conceptual, high-level explanation is called for. After all, being a car is likely a functional property for which we’ve had to
acquire information about reliable indicators. Even if you are unmoved by the example (there seems to be no property so rich that no theorist has argued it to be a part of visual phenomenology), at the very least, the cognitive model is a reasonable possibility: subtle features of visual input trigger a recognitional template which sets a certain cognitive module into motion. Absent independent reason to think otherwise, the mere discovery that control of $x$-detection by very subtle visual features seems neutral with respect to whether or not $x$ is perceptually represented.

The second line of argument proceeds from the way in which the effects discovered in the mentioned experiments were “irresistible”. Gao and Scholl write:

“[T]he data reported in these studies are all measures of visuomotor performance rather than explicit reports or ratings. This is an important distinction since overt decisions about what should and should not count as animacy can directly influence reports and ratings… but have no way to directly influence visual performance of the type studied here.” (ibid.)

Gao and Scholl illustrate this by revisiting the experiment above in which participants were tasked with avoiding a wolf that moved randomly a fixed percentage of the time. Even when explicitly informed about the details of the wolf’s movement, the participants’ performance often did not reflect this information. They write:

“[S]ubjects knew all about the random-motion interruptions, and had every incentive to discount them – but they could not do so in all cases.” (ibid.)

However, first, even if it’s true that animacy detection is immune to “overt decision” or, as they put it in the first line of argument, “conscious decision”, it still doesn’t follow that animacy detection is perceptual. It is evidence of a certain kind of modularity, but as I’ve mentioned above, it seems as though it’s at least possible for a module to output a non-perceptual state.
But putting that aside, the “irresistibility” Gao and Scholl claim here isn’t quite as cut and dried as, say, the way that seeing Muller-Lyer lines as differing in length is “irresistible” even though you know they’re of equal lengths. In Gao and Scholl’s case above, for example, there’s quite a bit that might get in the way in between being told “the wolf will turn no more than 20 degrees away from you” and taking part in the experiment. For example, as a subject, I might just have a difficult time translating that information into a visual expectation, or I might just have a difficult time translating that information into an advantageous strategy in the avoid-the-wolf game. (Should I scan the screen looking for objects moving away from me in order to mark them off as "safe"? Should I maintain a highly attentive vigil in the area immediately surrounding my disc and move away from anything I see moving toward me?). Thus, even if animacy detection were a matter of judgment, the new information might still face barriers to improving task performance. This kind of “irresistibility”, then, is also insufficient for distinguishing between perceptual animacy detection and cognitive animacy detection.

Last, Scholl and Gao argue from what they call “rich interaction with other visual processes”. They mention yet another condition which they call the “waverling wolf” condition in which participants view a display with three white discs. Two of the discs behaved as “sheep”, and moved randomly, while the third was a “wolf” which moved toward one of the two sheep at all times, switching which one it pursued every 1.2 seconds. By periodically marking the discs with probes and measuring probe detection rates at various stages of the display, Gao and Scholl gathered data suggesting that participants preferentially attended to the “wolf” even though the wolf was not identified to the
participants in any way and also preferentially attended to the “chased” sheep, whichever one that was at the time.

According to Gao and Scholl this shows that the discs were being preferentially attended to as *chaser and chassee*, considering the movement algorithm for the sheep was the same and so there was little difference in their low-level movement. The idea, then, is that this supports animacy perception in that this kind of control over another visual process is of a kind that is difficult to explain with any sort of judgment: since the wolf switched targets about every second, the allocation of participants’ attention was shifted consistently with “chasee” status in under a second, which, they hold, is too fast to be plausible if the animacy detection were a matter of judgment.

I find this a bit puzzling, since there seems to be a rather clear potential low-level explanation for this effect. While it's true, as Gao and Scholl say, that the movement algorithm for the two sheep was the same, I see no reason that the low-level theorist need restrict herself to properties of the "sheep" alone in explaining what might determine which sheep we attend to. Being in the trajectory of a moving object certainly seems to me to be a low-level feature (certainly lower than something as grand as the property of “animate”) of a display. Why couldn’t it be this, rather than animacy, that explains the preferential attending in the case of the two “sheep”?  

Further, something worth noting is that the line of argument above commits one to denying certain kinds of robust "cognitive penetration" effects, which roughly consist of various kinds of cognitive state having a direct effect on perceptual content. Cognitive penetration will be discussed in depth in a subsequent chapter of the present work, and detailed discussion can be found there. Here, however, we might just consider that the more
evidence we have that there are robust, quick cognitive penetration effects, as some argue there are, the less plausible it will be to suggest that if something is shown to be in robust, quick interaction with perceptual processes can, it must be perceptual as well.

In conclusion, though the work mentioned amounts to a number of impressive insights, Gao and Scholl haven’t shown that animacy detection is perceptual, and so haven’t shown that the property of animacy figures into visual phenomenology. Gao and Scholl are right to begin by ruling out cognitive explanations, but they fall short of this for the reasons mentioned above. So let us now move on to the case of faces, where there is a good case for a conscious perceptual effect, and the real point of contention is whether or not this effect is best explained by invoking rich content.

Case 3: Faces

Ned Block, seems to endorse the spirit of Siegel’s phenomenal contrast method, but insists that it cannot be applied on the basis of “armchair” considerations alone. This motivates Block to put together just the kind of case for the rich content it is the aim of the present work to evaluate.

Block’s argument comes in three parts all of which are based around perceptual adaptation effects. Most of us are familiar with such effects in some form or another already. When one stares motionless for a long time at a display with continuous motion, for example, and then quickly looks at a motionless scene, the motionless scene appears to be moving in the opposite direction.
There are a number of similar well-documented adaptation effects, and these effects exhibit similar properties to the ones advertised by Gao and Scholl for their chase-detection effects. The adaptation effects vary tightly with subtle details of the visual display. These effects persist even in cases where the subjects have reason to resist them, and the effects interact with other visual mechanisms like visual attention.

The key difference in the bodies of evidence, however, in virtue of which we have better reason to think that these adaptation effects are genuinely perceptual effects is that, unlike in the animacy cases, there is evidence that some of these effects are under the control of subtle visual details even when the subjects explicitly misjudge those details. Block cites, toward the end of Block 2014, the work of Schwierdzik, Ruff et al (2013), in which participants were asked to judge the tilt of a grid while staring at the grid for a long period, and then quickly shown a multistable grid (a grid of ambiguous tilt) and asked to estimate the tilt. The part of the result critical for our purposes is that even when the estimations of tilt of the first grid came apart from estimations of the second, the adaptation effect remained under the control of the actual tilt of the grid, and wasn’t disturbed by what the subjects thought they saw when they misestimated the tilt of the initial grid (Block 2014).

Note the way in which this is different from the Scholl and Gao result. In that case, participants were simply unable to (or reasonably presumed to be unable to) make conscious identifications of the controlling properties of the display at all. But in this case, we have subjects reliably transitioning from the presence of a certain visual detail to patterns of behavior despite an explicit incorrect judgement about that visual detail. The
result we find is that subjects can deny all they like that the grid is at a 20-degree tilt – the effect will still be the 20-degree effect.

Given this, let us move on to consider the part of Block’s argument most germane to the topic of this chapter. Block argues that just as the adaptation effect above makes it reasonable to accept that low-level properties like slant figure into conscious visual content, there are also high-level adaptation effects that give us reason to think that the involved high-level properties figure into conscious visual content. Block cites Butler, Oruc et al (2008) which details several variations on an experimental paradigm which is of the following general form: participants fixate for a long period on an angry-looking face and then switch to an emotionally neutral face, and then are prompted to rate the expression for anger. Subjects evaluate the face as less angry the more angry the habituating face looked (see below – try looking only at the picture on either the left or right and then only looking at the picture in the center), and interestingly, this effect can be found for a number of different high-level, social-property features like race and gender.

The argument I take Block to be making here involves generalizing from other cases of adaptation to this one. We all agree that certain properties figure into conscious visual content: properties like size, shape and color. These uncontrover-
sial properties all have in common that they are each subject to an adaptation effect specific to that property. So, if we find that there exists an adaptation effect specific to a high-level property, we should take this as at least some reason to think that this high-level property also figures into conscious visual content, thus militating in favor of the rich content view.

But how do we know this isn’t the result of low-level adaptation? That is, how do we know subjects weren’t just adapting to the low-level properties before them while looking at the angry face and making judgments on the basis of their altered, low-level phenomenology upon switching to the neutral face? Block writes:

"You can vary the low-level properties and so long as the emotional expressions are kept constant, adaptation obtains, though diminished somewhat, suggesting that face perception utilizes both low and high level attributives. It would take baroque congeries of low level properties to explain this fact." (Block 2014)

The idea here is that there is experimental evidence that the effect is diminished, but not eliminated, when the low-level features are varied but the high-level features are preserved. If the adaptation effect in the face cases were properly explained by adaptation to low-level facial features alone, the effect should disappear when the low-level features are properly controlled for.

In outline, this argumentative strategy is good. The problem is that, at least in the research he cites, the variation in low-level properties doesn’t exclude the low-level explanation. In Butler, Ocur et al (2008), for example, one of the ways the low-level
features was varied was to have subjects undergo an adaptation task which included conditions in which the facial features were translationally scrambled:

\[ \text{Intact face} \quad \text{Scrambled face} \]


As predicted, participants in the scrambled condition showed a much weaker adaptation effect. However, this is what the low-level adaptation theorist would likely predict also. Facial identification and emotional expression identification are widely held to involve at least two kinds of information: information specific to the shape and orientation of each individual feature, and information about the relative location of the features (see, for example, Fox and Barton 2007). While the information about the relative position of the features was, indeed varied in the Butler and Oruc et al experiment mentioned, information specific to each individual feature was not. Given this, we would expect just the result obtained: some of the low-level adaptation eventuating in face categorizing behavior was eliminated, but some was not, and a weaker effect was produced. This is true in general of the cases in which this kind of result appears in the face perception literature.

Potentially addressing this concern, however, is the last of the experiments Block mentions, which is based on shape adaptation. Susilo and McKone et al (2010) point out that we already have robust evidence of an “elongation” adaptation effect: staring for a long period at a very tall, thin oval causes subjects to misjudge a circle as short and wide. Given this, Susilo and McKone et al set up an adaptation experiment which involved gazing
at and estimating either a capital “T” shape of various lengths or an image of a face with the “T” shape formed by the eyes, nose, and mouth similarly varied (see below).


Subjects were put in an adaptation paradigm in all permutations of the above categories (they were adapted to a long T and asked to make judgments about an average-face, adapted to a long face and asked to make judgments about an average T, etc.), and the rather interesting result is that subjects showed only a weak adaptation effect in the T-to-Face and Face-to-T conditions, but did show a strong adaptation effect in the T-to-T and Face-to-Face conditions, which suggests the presence of face-specific adaptation since, if the adaptation effects involved were strictly shape-based, we might expect the same degree of
adaptation effect regardless of whether or not the images involved were Ts or faces - a low level shape adaptation mechanism shouldn't care whether or not it's adapting to a T, or adapting to the T-shape formed by a face's eyes, nose, and mouth.

There are, however, reasons to doubt the potential impact these results might have on the rich content view, intrinsically interesting as they are. For one, one might take issue with drawing such a conclusion from what is, in an important way, a negative result which might be explained in other ways. For example, one might wonder whether or not it really is the case that we should expect that looking at the stretched faces shown above would trigger a robust shape-adaptation effect in the first place. True, the eyes, noses, and mouths of the faces form a sort of T-shape, but the image itself has an ovular shape which, importantly, is the same size in each trial. If this is what subjects were adapting to, they might not show any differing behavior in the different face-to-T or T-to-face trials. This might give the false impression that no shape adaptation effect is present since the experimenters believed they were altering the relevant face shape in each trial.

Facial images (as well as actual faces) contain an awful lot of shapes. I know of no reason to think that any particular set is relevant for triggering low level shape adaptation. Presumably the matter isn’t just retinotopic the way, say, color or motion adaptation might be, since other examples of the shape adaptation effect involve different bounded shapes of differing sizes, subtending different angles in your visual field. But until such an account of shape adaptation is provided, the result the experimenters obtained might be just what the low-level theorist would expect. In the face-to-T and T-to-face conditions, the adaptation is weaker because whatever low-level adaptation is being triggered isn't transferring well between the two images' very different shapes. On the contrary, however
it is that shape adaptation works, in the T-to-T and face-to-face conditions all the shapes are roughly the same, and so we should expect a more robust effect.

A good starting response to this worry is to cite the second part of the Susilo et al experiment in which the same experiment was run but with the faces and Ts inverted. In this condition, adaptation effects in all four permutations were present. The idea here is that an upside-down face is no longer seen "as a face", and is seen, rather, as a batch of traditional, low-level properties, and so we have ordinary low-level adaptation effects kicking in equally for all permutations, unlike in the previous scenario in which a special face-specific adaptation effect caused differing levels of adaptation. As Block admits, no one experiment can rule out the above sort of concern, but this is at least suggestive, so let us put that concern aside and assume the experiment does show that there is a face-specific adaptation effect. It still seems unclear that this is the kind of adaptation effect that militates in favor of the rich content thesis. The Butler, Oruc, et al experiment, had the potential to confirm the rich content view, as the claim was that there was evidence that looking for a long time at an angry face caused one to have a visual experience as of less of the property “anger” than was actually present when looking at a subsequent face. But even if we grant that the Face/T experiment shows what its authors intend, this might only show that there is a face specific low-level adaptation effect. The experimenters themselves only argue that there is a shape-based effect specific to faces. All we would need to explain this, however, would be a low-level recognitional template playing a role in the adaptation effect. Perhaps, for example, (seemingly) face-specific adaptation processes become active only when we

---

10 Though I’m not sure how Block would want to interpret the fact that T-to-T adaptation was also stronger than face-to-T or T-to-face adaptation, since this seems at odds with the present explanation, I won’t pursue this here.
see certain families of eye-shape and mouth-shape meeting certain relative positioning criteria (the eyes above the mouth, for one).

I take it some of the difficulty here might be that it is somewhat unclear what Block would take to be a “high” enough high-level property to count as representing faces (or their emotional expressions) in visual experience. If face-shape-properties and spatial relations among them are enough to be called “face properties”, then I’m perfectly happy to agree with Block in allowing that there is evidence that these figure into visual phenomenology. However, this seems like it is, at best, a limit case for the rich content view, quite distant from the claims we’ve seen elsewhere. Siegel, in fact, regards the potential representation of pine tree shape properties as a threat to the claim that the property of being a pine tree is represented in visual experience, rather than vindication of it. This kind of property, at the very least, would be a less theoretically interesting addition to our conscious perceptual content.

Block closes with a brief argument against the foregoing experiments’ being the result of some sort of mere judgment rather than genuine perceptual adaptation. He argues that in the case of multi-stable percepts (such as ambiguous figure-ground stimuli like the famous vase/two-faces type of image) there is switching that occurs due to perceptual adaptation (see Alais, O’Shea et al 2010) which results in three key properties. In perceptual multi-stability: 1) the two percepts aren’t simultaneously present, 2) the percepts are transitory – one will eventually replace the other, and 3) the duration of each is not a function of time previously spent in either one. Block writes:

“Other things equal, if there were conceptual adaptation one would expect conceptual-without-perceptual adaptation – e.g. conceptually ambiguous situations would show alternations. For example, a morally ambiguous situation might lead
to alternating interpretations that show the same three fundamental properties. No such phenomena have been reported to my knowledge.” (Block 2014)

This seems to me to be too strong a requirement for those wishing to propose cognitive explanations. Unless the explanation is specifically invoking one, a cognitive explanation of the experiments above wouldn’t need to include a “conceptual adaptation effect” of the kind Block mentions to undermine the claim that a perceptual adaptation effect is involved. All that is needed in each case is an account that sufficiently explains the responses the participants in fact gave – for example, by invoking a response bias of the kind Gao and Scholl’s paradigm were designed to avoid (“I know these features are supposed to be good indicators of an angry face, and I notice they’re present, so I judge that I’m seeing an angry face.”). The theorist suggesting the alternative explanation needn’t, in addition, show that this bias is the result of a general conceptual adaptation mechanism bearing properties similar to our perceptual adaptation mechanism.11

I have already discussed why one might independently think that genuinely perceptual adaptation effects do exist, and it seems to me that they are the right thing to look for if one is looking to confirm the rich content view. However, I doubt that this kind of evidence deserves any general protection from alternative non-perceptual explanation and is, instead, something that should be evaluated case-by-case, since it is largely a matter of the difficulty of the case and the cleverness and skill of each experiment’s designers. But I won’t discuss this further here, since I have already granted that there is a genuine perceptual effect in each of Block’s cited cases.

11 And, further, there is some precedent for suggestive effects – Javadi and Wee found that when trained, not on faces, but on traditionally “masculine” objects like motorcycles and footballs, participants tended to rate a neutral-gendered face as looking more feminine (Javadi and Wee 2012).
In sum, Block’s argument does seem to reveal genuine, surprising perceptual effects, and seeking out evidence of adaptation does, indeed, seem like a potentially fruitful way of determining whether or not various content types figure into visual experience. In the above cases, however, adequate low-level explanations are available, undercutting the motivation to accept the rich content view.

Case 4: Causation

Up to this point, we have considered cases of potential rich content which all have at least one similar problem (in addition to there being others they do not all share): in each case the high-level property under consideration supervenes on low-level properties subjects were quite plausibly visually experiencing. Even if it is the case, for example, that participants couldn’t tell us that the wolf was chasing them 30% of the time, the participants could certainly have told us about each of the specific movements they witnessed of which the property “chasing-30%-of-the-time” is a function.

In the case of mechanical causation, however, there appears to be an example in which this isn’t true. Stephen Butterfill, in discussing the work of Albert Michotte, points specifically to part of his work involving “launchings”, in which a circle is depicted as approaching another circle and stopping on contact at which point the other circle is depicted as moving away at the same speed (Butterfill 2009, Michotte 1963). In the experiment, the variable is the amount of time the display pauses in between the time of contact (when the first circle reaches the second and freezes) and the time the second circle begins moving away. The display varied at intervals of 14 ms of pause, and at each interval
participants were instructed to determine whether or not the first circle caused the second to move. Unsurprisingly, the greater the pause, the less likely participants were to report having seen a causal event.

Thus far, this result is similar to other empirical work discussed here, and might be vulnerable to similar responses. What is of particular interest in this case, however, is what Michotte found while collecting additional information from participants. Using the same sort of displays varying in the same increments of pause, Michotte asked participants various questions comparing multiple displays. Participants who had just seen two launching event displays differing only in 14 ms' worth of pause (for example, a subject would be shown a launching event with a 28 ms-pause and then one with a 42 ms pause) would, when prompted, report that the displays were identical - that is, when shown a launching event with a certain length of pause, and then shown a different display with a 14 ms longer pause than the first, the participants failed to report any difference between the two events.

However, there was one pair of cases for which this general rule didn't seem to hold. Subjects were able to identify that a pair of displays differing only in 14 ms' worth of pause were, indeed, different as long as the change in pause between the two displays crossed the 84ms boundary. So, for example, although subjects who were shown a display with a 28 ms pause and then a display with a 42 ms pause reported that the two events they'd just seen were identical, subjects who were shown a display with an 80 ms pause and then shown one with a 94 ms pause were likely to report that the displays looked different somehow. This is rather puzzling - why should certain 14 ms pause changes be reported, but not others?
Butterfill argues that we ought to take this as evidence for the rich content view because it is evidence that we are experiencing displays with a less-than-84-ms pause as instances of mechanical causation. He likens it to the well-accepted model of speech perception, on which we have reason to say we perceive phonemes because we perceive them categorically – a speech sound is either heard as a /b/ or a /p/, but nothing in-between. According to Butterfill, Michotte’s data suggests the same kind of conclusion. There appears to be a perceptual category boundary at or near an 84 ms pause in a launching display, since participants can detect a 14 ms change in pause only if it crosses the 84 ms boundary, and this, along with participants’ reports gives us reason to think this boundary corresponds to the difference between visually experiencing events as causally related and not.

In a case, the high-level explanation would seem to have a genuine explanatory advantage in a way unlike the other cases considered in this chapter. According to the low-level theorist, we only have a visual experience as of two episodes of motion separated by a pause. Thus, the low-level theorist has to come up with some sort of additional story explaining why the 84 ms boundary should be privileged, and might cause participants to claim the displays "look different" only when on opposite sides of the 84 ms boundary. The high-level theorist, on the other hand, can explain this easily. The participants, the high-level theorist can hold, do not visually represent intervals of pause at a fine grain (at least, not to the millisecond), and this is why participants insist displays differing in pause by 14 ms look identical. They do, however, visually represent displays with less than an 84 ms pause as causal, and the others as non-causal, and this is why displays crossing that boundary are reported as looking different.
Unfortunately, there are a few problems. As Lance Rips points out, the data are not as clear as one might hope (Rips 2011). For there to be good evidence of a category boundary there should be a very sharp change in the difference-detection rate before and after the 84 ms interval which we don’t see in Michotte’s data. Michotte seems to have agreed on this. Rips quotes Michotte:

“In the present case, as in all cases involving perception of form, the differences are not clear cut. Just as it is possible to have the impression of a shape that is more round or less round and of one that is more square or less square, so it is possible to have the impression of a Launching Effect, or in general, a causal influence, which is more pronounced or less pronounced, more clear or less clear, “better” or “poorer”. These are differences in degree, and these correspond to definite stimulus-conditions and to the size of the time-intervals in particular.” (Michotte 1962, p. 95, via Rips 2011)

But let us put this aside - even if the data showed an enormous shift at the 84 ms boundary, the low-level explanation wouldn't be ruled out. For example, Butterfill's argument from categorical perception is just as much of an argument for coarse-grained pause-perception as it is for cause-perception: the only thing Butterfill's argument rules out, if successful, is that we represent the display at a very high temporal resolution, but this isn't the only viable low-level explanation of the participants' behavior. Perhaps, for example, the 84 ms boundary change in behavior ought to be explained by participants' having seen displays with a shorter pause as not containing a pause at all (as a display with continuous motion). It's certainly consistent with the low-level theorist's aims to hold that participants are experiencing this property. We might then account for the general way in which participants can detect large pause changes but not very small ones with a different mechanism entirely - perhaps a non-perceptual mechanism for roughly tracking the passage of time which the participants are employing once they see a display as including a pause.
Further, even if the low-level theorist is restricted to attributing to the participants representations of displays as containing pauses of lengths of \( n \) milliseconds, the results of Michotte's work still don't show that participants weren't experiencing the displays in this way. They could have been representing, in visual experience, two displays that differed in 14 ms' worth of pause without anywhere representing \textit{that} they differed in pause (other than at the 84 ms boundary). Put a bit more roughly, if Butterfill is worried about the compatibility of a low-level explanation and the participants' having visual experiences with the same content, one response is that it might be that the participants weren't having visual experiences with the same content, and they, in some sense, just didn't \textit{notice}. One possible way to flesh this suggestion out is to hold that we do represent the pauses down to the millisecond in visual experience, and we represent them \textit{as pauses}, and nothing fancier than that. However, we do have a \textit{perceptually sensitive cognitive module} which takes as input our perceptual experience and outputs judgments about causation (or, at least, some sort of cognitive representation that eventuates in such a judgment). As before, since we can explain the data with low-level perception and perceptually sensitive cognitive module, Michotte's experimental work doesn't push in favor of mechanical causation figuring into conscious visual experience.

\textbf{Conclusion}

In this section, I considered the claim that visual phenomenology includes more than just "traditional" visual properties like size, shape, color, and motion, and ultimately concluded
that, despite considerable recent interest and development, in most cases we still lack reason to think the rich content view is true.

I considered Susanna Siegel's "phenomenal contrast" method of investigation, but found reason to doubt it would allow for better armchair theorizing than ordinary, unconstrained inference to the best explanation, and proceeded to consider some of the best cases for rich content on an individual basis. Fish's argument from attention was argued to be a non-mandatory interpretation of the cited empirical work by Li et al. Gao and Scholl's work on animacy perception was considered, but turned out to be insufficient to show that the effects they found were genuinely perceptual, let alone a part of conscious visual experience. Block's argument from face perception was considered, and while aspects of face perception do seem to be genuinely perceptual, genuine high-level attributives are not necessary to explain the empirical results Block invokes. Finally, we considered the visual experience of mechanical causation by considering the work of Albert Michotte. I argued the data itself has been merely suggestive, at best, and that even if it were better, this would still not rule out a low-level explanation involving a perceptually sensitive module.
Chapter III: Cognitive Penetration

Naively, we take our visual experience to be an excellent provider of justification for many of the beliefs we form directly on its basis. If I want to know what color my neighbor's car is, I take a look out the window, I have a visual experience as of certain blue shapes, and, as long as I haven't suffered a head injury or taken any experience-altering substances lately, it's strongly intuitive that I would be justified in forming a belief that my neighbor's car is blue on the basis of that experience.

In this chapter, however, I consider issues surrounding both hypothetical and (purportedly) non-hypothetical cases of "cognitive penetration": cases in which our perceptual experience may have been influenced by our own antecedent cognitive states, before reaching conscious awareness. In cases like these, our epistemic intuitions can become much less clear. What if the young man in front of me looks to be holding a weapon only because my visual experience is being affected by a terrible racial prejudice of mine? With the addition of a questionable etiology to the story, many of us have the intuition that something is epistemically amiss in the situation, and our intuitions about what justifying power my visual experience has become cloudy.

In the present chapter I discuss cognitive penetration cases and what they might mean for several theories of perceptual justification. First, I discuss the general epistemic problem posed by cognitive penetration cases as presented in The Rationality of
I then proceed to discuss several responses to the problem. First, I discuss the problem in connection with internalist, "dogmatist" views, and discuss a response by Michael Huemer, who argues that cognitive penetration cases should not be treated differently than any other case of perceptual justification - that is, even in cases of penetrated experience like the one above, the experience has as much justificatory power as it would in an unpenetrated, otherwise-identical scenario. Next, I turn to discuss cognitive penetration cases in connection with externalist theories of justification, including responses from Jack Lyons (2011, 2016), which predict a loss of justificatory power in certain kinds of penetration cases. In the remainder of the chapter, I discuss three responses to this externalist account of cognitive penetration cases. Each one is an attempt to work out an internalist account of cognitive penetration cases that can, nonetheless, respect the intuition that, in the intuitively bad cases, the penetrated experience loses justificatory power. The first is an account based on the claim that experiences can have bad etiologies in the same way beliefs can. The second is an account based on a notion of experience-generating inference ("quasi-inference") which makes experiences epistemically good or bad depending on whether or not an inference to a belief with the same content as that experience would have been a good inference for the subject. The last is Siegel's most recent account, according to which perceptual experiences can, themselves, be rational and irrational and experiences redound on the rational standing of the experiencer. I argue that although the three considered accounts are the strongest proposals for a middle-ground internalism, each faces serious difficulties, and there is reason to think that the general approach is under-motivated in

85
the first place. I conclude that views that can accommodate both intuitions have a theoretical advantage.

The Cognitive Penetration Problem

To see the particular problem posed by cognitive penetration cases, let us begin by considering the main expository example in Siegel (forthcoming). We are asked to imagine the case of Vivek, the vain performer. Vivek has a very high opinion of his own abilities as a performer. As he performs, he looks out upon the audience and sees a sea of faces ranging from pleased to neutral in expression. Remarkably, Siegel writes, nobody ever looks displeased. Vivek considers this evidence, and on this basis, strengthens his belief in his own abilities. We should get the general impression that, at least, Vivek's belief formation/strengthening process includes some manner of serious epistemic flaw in virtue of the influence Vivek's vanity has had.

This story, however, is full of (deliberately) ambiguous terms, and we might note that there are many possible ways of filling out the details of this story. Vivek might, for example, be subject to something like a bias in perceptual judgment rather than having different perceptual phenomenology. That is, perhaps Vivek has the same visual phenomenology you or I would have in his position, but his vanity causes him to judge audience faces to be displeased only at extreme levels of visual evidence. Most would agree that there is something epistemically wrong here, but this would not pose a new problem peculiar to the kind of case under consideration. There is little temptation to think beliefs formed in this way would be justified, as jumping to conclusions (the way
Vivek has done here, motivated by vanity) is a paradigm way of generating an ill-founded belief. Vivek's situation would, as Siegel puts it, be a "perverse self-perpetuation [which] epistemically compromises the subject." (Siegel 2017, p.16)

It is only when Vivek's vanity strikes very deeply, affecting the content of his visual experience that a distinctive philosophical problem arises. That is, the cases which pose the special problem under consideration here are the ones in which Vivek's vanity injects influence into his visual phenomenology and Vivek has a visual experience as of a generally pleased crowd, whereas, if he had had more accurate beliefs about his talents, he would have had a different visual experience.

Unlike the scenario in which Vivek is merely making poor judgments due to vanity, the scenario in which vanity is affecting Vivek's visual experience generates a pair of opposing pressures. Siegel illustrates this with a different case:

"Suppose Jill fearfully suspects that Jack is angry at her, and this makes her experience his face as expressing anger. Her experience is a farce - Jack's face is neutral, and Jill could see that just as plainly as others can, if her suspicions weren't interfering with her perception." (ibid)

In this scenario, unlike the bad-judgment scenario, it can seem as though there's pressure to say that Jill is behaving reasonably in forming her beliefs. After all, as Siegel puts it, "if Jack really does look angry to Jill when she sees him, and she has no indication that the experience is misleading, then what else could Jill reasonably believe about his emotional state, other than that he is angry?" (ibid)

On the other hand, we should note the way in which the case of belief resulting from Jill's fearful experience can still seem to result in an epistemically defective belief. Siegel puts it the following way:
"This situation seems epistemically pernicious... if behind the scenes, the penetrating states are stacking the tribunal of experience in their own favor, then while experience will seem to let you check your beliefs against the world... really you'll just be checking your beliefs against your beliefs. The tribunal will be corrupted. On the face of it, epistemic elevation in such a circumstance seems illicit." (Siegel 2011)

One might, to expand on this worry, imagine Jill repeats this process over and over: she starts with an unfounded suspicion, looks at Jack, forms a belief, then looks at him again. Now he looks even more clearly angry in step with the formation of the belief. Then, on the basis of that experience, Jill strengthens her belief, and starts all over again. Jill, through a laundering process, gets herself increasingly better grounds for the belief that Jack is angry every time she looks. Surely, one might think, this is somehow an epistemically poor route to belief; justification can't come from nowhere. Jill is "bootstrapping": her belief would seem to be mysteriously (and implausibly) pulling itself up to higher and higher degrees of justification merely by tugging at its own feet. Siegel compares the situation to a gossip circle, in which "Jill tells Jack that p, and Jack believes her but quickly forgets that she's the source of his belief, then shortly afterward Jack tells Jill that p"12, noting that it would be silly to think that Jill would be reasonable in strengthening her belief on this basis.

Note also, that what makes these cases philosophically interesting in a different way from standard cases of illusion and hallucination is that when such cases are discussed in the philosophy of perception, the illusions and hallucinations typically don't arise because of a content-based connection with other mental states the agent has and often don't include anything about the etiology of the illusion or hallucination at all. This,

12 Let us assume, also, that Jill has forgotten she was the one who told Jack that p.
at least at face value, also changes the intuitions we have about the resulting belief. Imagine a case where Vivek's visual experience has nothing to do with his vanity - he was just struck with a sudden, brute, unlucky hallucination like a lightning bolt, which just happened to amount to a visual experience as of a pleased crowd. The intuitive tension seems to dissipate. It's now at least more intuitively comfortable to consider that Vivek's belief about the crowd might be justified. The robust intuitive tension seems specific to cognitive penetration cases.

Siegel invites us to consider a variety of examples which seem to generate this intuitive tension, including a (likely) fictional case of preformationists who seemed to see embryos when looking at sperm cells under early microscopes and a (purportedly) non-fictional case involving experimental subjects who, it is supposed, looked at grayscale images of bananas and had an experience as of a yellowish banana because of the belief that bananas are typically yellow. In each of these cases, the same two conflicting pressures arise: on one hand, one doesn't want to license bootstrapping, but on the other, if the banana really looks yellow, what else could it be appropriate for the subject to do?

Siegel mentions one additional sort of case at this point, however, which expands her discussion a bit beyond cases of cognitive penetration proper. This is a case in which the "penetrating" state affects patterns of attention, rather than directly altering the content of perceptual experience. Siegel invites us to imagine being in the middle of reviewing applicants in order to hire one for a position. As Siegel notes, it is well-confirmed that hiring committees respond differently to applications from socially

---

13 Siegel calls these "zap cases". See Siegel (2013a)
14 Siegel's intent is to cover a wider range of phenomena which she refers to as "perceptual farce". Cognitive penetration cases are one kind of perceptual farce.
stigmatized and socially dominant groups. We might easily imagine a case in which this occurs because, as Siegel writes,

"the features that detract from the stigmatized applicants' candidacy are selected for further processing, while features that support it could be anti-selected. A fuller picture of the candidate would reveal counterbalancing strengths, but the strengths in the application are simply not registered to begin with. From the evaluator's point of view, the stigmatized applicants simply have little or nothing to recommend them." (Siegel 2017)

Siegel claims the same two pressures arise in cases like these.

As Siegel notes, cases like these do present strong reasons to be interested in mechanisms of belief that wind up having this sort of self-reinforcing character, as these might be the mechanisms behind things we care about for independent (in this case, moral) reasons. According to Siegel, this sort of attentional effect might be a mechanism through which unfortunate, oppressive social groupings are perpetuated. (One might groundlessly believe a certain class of persons to be ill-suited for a certain kind of difficult work, that belief might influence one's perceptual experience, through attention or otherwise, and one might take that experience to confirm the claim that that class of persons is ill-suited to that work). However, the moral dimension adds undesirable intuitive complications. For example, even if one might be epistemically justified forming the regrettable belief that the applicant is underqualified, it may be the case that in a situation like that one is morally obligated to conduct more thorough research due to the great harm that would result if one is wrong. Because these kinds of potential normative interaction are too complicated to discuss here, I will restrict discussion to certain kinds of examples in order to try to minimize potential normative muddying.

A Traditional Internalist Response
The two pressures mentioned in the above cases present a challenge to what I will call “Dogmatist” views of justification, following Jim Pryor (Michael Huemer’s phenomenal conservatism is similar, though is a slightly broader term). According to the family of views under consideration, if one has a perceptual experience as of p then, absent defeaters, one has direct prima facie justification for believing p without needing anything else to support it (such as beliefs about the workings of one’s perceptual system, beliefs about norms of rationality, etc). That is, perceptual experience as of p is, all by itself, enough to justify believing p all by itself (absent defeaters). Whatever is beyond the agent's ken, on traditional views of this sort, is unrelated to justification (on varyingly loose understandings of what counts as within one's ken).

The dogmatist, therefore, will have to make the case that Vivek's resultant belief is justified, and explain away the intuition we have in cognitive penetration cases that something is epistemically amiss.

In light of the introduction above, which was focused on making clear the intuitive pull behind extending epistemic downgrade in Vivek-style cases of cognitive penetration, this might seem unintuitive, and this is precisely what Siegel argues (Siegel 2011, 2013a, 2017). But in response, we might consider a “subject’s point of view” style of response like the one found in Michael Huemer’s response to Siegel (Huemer 2013).

In discussing cognitive penetration, Siegel invites us to imagine that she is terrified of handguns, and that this fear penetrates into her visual experience when she opens the refrigerator late one night. Her fear makes her have a visual experience as of a handgun
in the fridge, but, she claims, it’s unintuitive to think she is justified in forming a belief that a gun is in her fridge on this basis.\(^\text{15}\)

In Huemer’s response, he invites us to think of the situation from the subject’s perspective: she has no beliefs or suspicion about the way her fears might influence her perceptual experiences. She just opens up the fridge one night and, lo, a gun! We should, he insists, think of it in the following way: if what the subject would be rational in doing is to withhold from forming the belief that there is a gun in the fridge, this is quite strange. What reason could she possibly give for this?

In a rather vivid illustration, Huemer asks that we imagine that there are eggs in Siegel’s refrigerator and she has a visual experience as of eggs in the refrigerator as well as of a gun in the refrigerator. The gun-experience is a hallucination resulting from fear-based cognitive penetration, while the egg-experience arises as an ordinary response to the eggs which are, in fact, in the fridge. Siegel is, as already mentioned, in no way in a position to know about any of this. The anti-dogmatist intuition, Huemer argues, would be that what it is rational for Siegel to do is to accept the claim that there are eggs in the refrigerator while rejecting or withholding from the claim that there is a gun in the refrigerator. Suppose Siegel does this. If asked why she accepted one proposition and not the other Siegel would have nothing to say. Huemer writes

"She might thus report her state of mind in some such fashion as the following: 'I accept [E, the claim that there are eggs in the fridge,] on the basis of my visual experience. [G, the claim that there is a gun in the fridge,] also appears to be equally well-supported by my visual experience, and I have no reason for thinking the experience representing G to be any less reliable, nor epistemically inferior in any manner whatsoever, to the experience representing E. Nor have I any other grounds for doubting G. Nevertheless, while I accept E, I refuse to accept G, for

\(^\text{15}\) For our purposes here, let us ignore that there may be a defeater here due to the extreme likelihood of a gun appearing in an ordinary person's refrigerator (hopefully, this would be enough to make most of us do a double-take).
Huemer continues by suggesting an explanation for the intuition that there has been an epistemic foul in stories like Vivek's. Though most agree on a strong intuition that something is epistemically amiss with Jill’s formation of the belief about Jack’s anger on the basis of the penetrated experience, perhaps we might explain this intuition as tracking the way in which cognitive penetration cases like Vivek’s amount to an obstacle to knowledge, rather than an intuition about justification specifically. That is, cognitive penetration cases may be accounted for in the same style as one's favorite account of certain Gettier cases: though, Jill has a justified true belief about Jack's anger, the questionable etiology of her experience puts knowledge out of her reach.

Siegel anticipates this suggestion, arguing that this can't be the right explanation of the downgrade intuition, since the intuition is that, in cognitive penetration cases like Jill's, we have the intuition that the penetrated experience somehow loses epistemic power whether or not Jack is, in fact, angry. She writes:

"Intuitively, Jill's fear-penetrated anger-experience puts her in a worse epistemic position than she would be, if she had [a non-problematically-penetrated] experience, all other things being equal. But if Jack really isn't angry, then Jill can't know on the basis of her anger-experience that he is, whether her experience is [problematically penetrated] or not." (Siegel 2013a)

If Jack is not actually angry, Siegel argues, knowledge is out of the question already. Yet the intuition that something is wrong persists, and so this intuition cannot be explained as tracking the fact that knowledge is beyond the agent’s reach. Against this, however, it seems as though one's downgrade intuition might be explained as tracking an obstacle to knowledge (perhaps the violation of a Gettier-type rule) even if it turns out Jack isn't
really angry: the false belief would merely amount to another obstacle to knowledge. To put it another way, assume that knowledge requires true, justified belief, as well as one’s favorite anti-Gettier factor. It’s plausible that our intuition that Jill is worse off with a penetrated experience even though the belief she forms on its basis is false is just the intuition that, in the penetrated case, Jill is failing to meet two of the necessary conditions for knowledge rather than one. Perhaps this might take the form of a modal intuition: Jill’s belief about Jack’s anger is justified and false, and we recognize that even if it had turned out to be true, this still wouldn't amount to knowledge for Jill despite her having a justified belief to that effect. If this is a plausible account of our downgrade intuition in the above case, then cases like these fail to undermine the aforementioned "obstacle to knowledge" type accounts.

To be clear, however, this line of defense for the dogmatist does not successfully dispel the intuition that there ought to be epistemic downgrade in cases like Vivek’s. I defend it here in order to defend a particular kind of account as a live option: one which makes room for both of the seemingly opposed intuitions to which cognitive penetration cases give rise by suggesting an additional epistemic dimension relevant to these cases along which the agent might be failing.

Again, my claim is that cognitive penetration cases result in two seemingly opposed intuitions, and neither is easily dispelled or explained away as theoretically insignificant or misguided. So, it is desirable that a theory respect both, if possible. I am, myself, attracted to the idea that our concept of justification may just have multiple components to it that are sometimes in tension (perhaps there are both accuracy-oriented
components and behavior-guiding components, for example). But for now, I mean only to bring out and defend both intuitions which ought to be respected in analysis.

**An Externalist Response**

Unlike dogmatists, however, externalists about justification can quite naturally respect the intuition that something has gone wrong in cases of cognitive penetration. It is available to the externalist to hold that, with respect to justification, Vivek's belief that his show is going pretty well *is* worse off due to his penetrating vanity because his penetrating vanity is decreasing his reliability. In the example, Vivek's vanity causes him to see the crowd as neutral-to-pleased, whether they have those expressions or not. Vivek's vanity has caused him to become perceptually insensitive to disapproving faces which makes the beliefs he forms on that basis of his perceptual experience prone to error - his vanity is making him form beliefs in a way less likely to track the way the world actually is.

As Lyons (2011) argues, an advantage of the reliabilist account is that it can divide cognitive penetration cases into intuitively "bad" and "good" cases, as sometimes cognitive penetration seems epistemically vicious, but sometimes it seems epistemically innocent, and the reliabilist account is in a good position to sort the cases well.

Consider two cases:

**Bad Snake:** Perceiver BS's expectation of snakes cognitively penetrates her perception by selectively shifting down the match criteria or threshold for her snake template. The "tuning curve" of this snake detector is thereby broadened, and a wider range of stimuli sill now trigger it, as the result of [cognitive penetration]. Suppose further that top-down consequences of snake detection alter low level visual experience of color, shape, etc.
**Good Snake**: The tuning curve of GS's snake detector remains highly selective, but the detector is, due to CP, queried early and often in the perceptual process, thus leading to increased - but non-biased - processing of snake-like stimuli. That is, the penetrating expectation influences which information gets selected for further processing (namely, the information potentially relevant to snakes) but not the results of this processing. (Lyons 2016)

Bad Snake is a case much like Vivek's or Jill's: BS has an expectation penetrating her visual experience in a way which affects what it takes for her visual experience to include snakehood in content, and which also alters the content involving low-level properties in her visual experience. Much like in the case of Vivek or Jill, we should have the intuition that something would be wrong with BS's forming beliefs on the basis of this experience. In Good Snake, however, our intuitions are different. Much like in certain cases of visual learning or some cases of cognitive penetration by expertise, it seems much more intuitive that the belief GS forms on the basis of her visual experience is justified.

Externalists, as suggested above, can explain this quite neatly. It is available to an externalist to hold that Bad Snake seems bad because the belief was formed on the basis of an unreliable process, and so isn’t justified. BS's expectation causes her to see things that are only roughly snake-like as a snake, and changes the low-level properties that figure into her visual experience into more snake-friendly configurations of low-level properties. Her perceptual experience will be flooded with false positives, and so she is not justified in forming related beliefs solely on that basis. GS, on the other hand, stands to be an excellent snake detector. Since her system isn't doing any "overwriting" of low level properties and isn't registering loose matches as snakes, but is instead doing more checks for snake-matches, we can expect that GS will miss fewer snakes in her field of view without also causing the number of false positives to explode. Because GS's perceptual system will do a pretty good job of accurately representing the world (with respect to snakes) due, in part,
to the penetrating expectation, we don't have the bad epistemic intuition in this case because the snake-beliefs GS is forming on the basis of her performed on the basis of a highly reliable mechanism.

In accounting for the intuition that there is something wrong in cognitive penetration cases, reliabilism seems well-suited to the task. It seems to do a good job of sorting cases that are "good" in a certain, intuitive way, from cases that are bad in that way. But we also saw that single cognitive penetration cases can give rise to two seemingly conflicting intuitions, and that this is true in many of the cases that Lyons is calling "bad" cases. In these "bad" cases, there does seem to be something wrong with Jill's belief about Jack, but, again, there is also an intuitive pressure to protest, "what else was she supposed to do?" That is, in "Bad Snake", if the subject's visual experience really represents the presence of a snake, or snake-like low-level properties, then, absent defeaters, how could it fail to be rational to believe a snake to be present?

Mentalism Part 1: Siegel (2013)

We have discussed two approaches to cognitive penetration cases each of which is well-equipped to handle one of the two seemingly opposed intuitions such cases arouse. Insofar as these intuitions really are opposed, each account should be concerned to dispel or explain away the intuition unfriendly to that account.

As discussed, this appears difficult to do. However, the two kinds of intuition are only incompatible insofar as they push in opposite directions over a single dimension of epistemic evaluation. It seems to me that, in the absence of a clearly theoretically superior
account, instead of pressuring us to find a way of explaining away one or the other intuition, the above considerations push us toward accepting a view according to which there are simply two dimensions of doxastic justification - one that answers to the internalist sort of intuition, and another that answers to the externalist sort of intuition.

John Greco, for example, has argued for a mixed view, incorporating both a notion of justification couched in terms of "epistemic responsibility", answering to the internalist sort of intuition above, and a condition on knowledge that includes "objective reliability" that must occur as a result of the subject's being epistemically responsible (Greco 2000). The specifics of the interactions between these factors are beyond the scope of the present discussion. The point here is merely that one need not take the two intuitions to be opposed, and admitting of two dimensions of doxastic evaluability can result in fruitful theorizing. I contend that unless a clearly theoretically superior account emerges, giving us reason to think we ought to explain away one or the other intuition, the intuitions that cognitive penetration cases stir up give us reason to consider a pluralistic view.

In what remains, I will consider three attempts to produce just such a theoretically superior account. I begin with Siegel (2011, 2013a), in which a third option is proposed: a mentalist view which is intended as a an internalist account which can, nonetheless, accommodate the intuition that, in cases like Vivek's, Vivek would not be justified in strengthening his belief in his abilities on the sole basis of his visual experience as of a neutral-to-happy crowd.

Unlike traditionalist internalisms like Huemer's, according to mentalist internalisms like Siegel's, justification isn't tied to any notion of what the agent has access to. (Recall the passage from Huemer, which made the case for the dogmatist intuition from the agent's
"point of view", couched in terms of what the agent is and is not in a position to appreciate). Instead, on the kind of mentalist account Siegel favors, justification merely has to supervene on the mental states of the agent, with no mandatory accessibility-related restrictions. Siegel proposes that, although a mentalist account of cognitive penetration cannot invoke considerations of reliability or accuracy, it nonetheless has the resources to predict epistemic downgrade in the bad cases of beliefs formed on the basis of cognitively penetrated experience, making it a competitor in the present dialectic for the explanatory spot currently held by forms of reliabilism.

Roughly, the view in Siegel (2011, 2013a) is that perceptual experiences, while not themselves rationally evaluable, do have rationally evaluable etiologies. If an experience has a bad etiology, according to Siegel, it suffers a downgrade in its ability to justify beliefs based on it, even though it is itself neither justified nor unjustified. Of course, what counts as a bad etiology must be specified in a mentalist-friendly (non-reliability-invoking) way, and so Siegel offers just such a framework: an experience's etiology is a bad one just in case a belief with a sufficiently similar etiology would be unjustified. She writes:

"[T]here is such a thing as an experience with irrational etiology, such etiologies epistemically downgrade experiences, and barring an assimilation of rationality to reliability, this source of epistemic downgrade is distinct from unreliability." (Siegel 2013a)

Siegel then lays out her account in the following way:

"[An experience with content C and etiology X has reduced justificatory power iff] a belief with content C and etiology X* would be doxastically unjustified, where the output of X* is a belief with no intervening experience, and X* has psychological elements sufficiently similar to X's." (Siegel 2013a)
The idea is that there are paradigm *doxastic* etiologies that clearly result in beliefs that do not have the power to justify other beliefs formed on their basis: wishful thinking, circular reasoning, jumping to conclusions and the like. Why shouldn’t we expect that *experiences* that arise through the same sorts of routes would suffer the same downgrade in justifying power? Siegel argues that we should.

Siegel’s mentalist analysis, then, of Vivek’s case is as follows: Vivek’s belief strengthening is unjustified in virtue of having engaged in a bad experience formation process. His experience was the result of "wishful seeing" (the perceptual version of wishful thinking), and so it has a reduced ability to justify beliefs formed solely on its basis. After all, he wouldn’t be justified in forming the belief solely on the basis of wishful *thinking* – wishful thinking processes do not afford prima facie justification.

Again, the claim is that if beliefs have reduced justificatory power as a result of a certain sort of etiology, then we ought to think that experiences with that same sort of etiology should as well. But, as Siegel is aware, this is only the case as long as we can find no relevant disanalogy between the belief cases and the experience cases.

Siegel attempts to provide some argument in support of the claim that we won't find such a disanalogy. She writes:

"When ill-founded beliefs generate more ill-founded beliefs (for instance when you base belief B2 on ill-founded belief B1), do they do so by virtue of any features distinctive of beliefs, that experiences do not share? If so, that would undermine the idea that checkered experiences generate ill-founded beliefs, by virtue of their etiology." (Siegel 2013a)

She then proceeds by examining what she takes to be the most plausible potential differences between experience and belief, arguing that none of them are relevant to the
way in which ill-founded beliefs (with a bad etiology) generate further ill-founded beliefs.

She considers several to which I will return at the end of the chapter in considering Siegel's newer view. Of particular interest at this stage, however, is the clear difference already noted above, which is that beliefs can, *themselves*, be irrational, while experiences on an orthodox view, cannot. In the case of belief, being a justified belief and having the power to justify other beliefs go together. That is at least some reason to worry about the claim that experiences with similar etiologies will have analogous differences in justifying power. Siegel, however, argues that a state’s being rationally evaluable cannot be necessary for the transmission of ill-foundedness, since beliefs sometimes do help transmit irrationality without themselves being irrational. She writes:

"Often beliefs are based on clusters of other beliefs, rather than on single beliefs, and not every belief in the basing cluster need be irrational, in order for the subsequent belief to end up that way. If I rationally believe that I'm going to New Jersey today and irrationally believe that in New Jersey it is raining locusts, I can end up with an irrational belief that I'll likely see some locusts when I get there, thanks in part to my rational belief about where I'm going. So a belief need not be irrational, in order for it to generate a subsequent ill-founded belief formed on its basis. Thus the fact that experiences are never irrational does not preclude them from doing the same." (Siegel 2013a)

Siegel’s argument, however, is aimed at the wrong kind of ability for the dialectic. It does seem as though rational beliefs can be *part* of the cluster of things on which an irrational belief is based, and insofar as beliefs can be based on experience, it does seem as though experiences can also be *part* of the basis for an irrational belief. Belief and experience do not differ under this description. Presumably, dogmatists would happily accept this. However, this isn’t specific enough to support Siegel’s main claim: we wanted to
determine whether or not an experience’s *ability to justify* can be affected by a bad etiology, and it’s not clear that the case above is specific enough to bear on this issue. For example, in a case of perfectly well-founded belief, I might rationally believe I’m going to New Jersey today and rationally believe that in New Jersey it’s raining water, resulting in the well-founded belief that I’ll likely see some rainwater today. Even though in this case the resulting belief is well-founded and in the locust case the resulting belief was ill-founded, it seems quite plausible that the rational belief that I’m going to New Jersey today, *itself*, is playing the same role with respect to justification in both cases: it doesn’t seem as though one can get a clear picture of the justifying powers of a belief just by looking at the justifying powers of a multi-member set of beliefs to which it belongs.

To extrapolate a bit, in the paradigm cognitive penetration cases, we have the intuition that there's some sort of epistemic failing in scenarios in which beliefs are based *solely* on a single perceptual experience. This means that, if experiences cannot be rational or irrational, then cognitive penetration cases are cases where the basis for the beliefs contains *no irrational states at all*. Given this, what Siegel should be seeking for support are cases where a non-irrational state is a *source* of epistemic badness that "infects" a belief formed on a basis in which that state is included. A doxastic scenario that fits this bill would be something like a case in which there is a set of beliefs *none* of the members of which are ill-formed which, nonetheless, "infects" the resulting belief which is based solely on that set in an otherwise responsible way. It seems, however, far from clear that this can be found.

Putting this looming difficulty aside, the account also faces some more immediate threats. Recall that what Siegel aims to present is a mentalist account on which we can
predict epistemic downgrade in bad cases of cognitive penetration without invoking externalist considerations like reliability. Siegel mainly focuses on similarities and dissimilarities between beliefs and experiences generally. But recall that the critical part of the account is, roughly, that experiences have a reduced ability to justify beliefs formed on their basis just in case a belief with an etiology having "sufficiently similar psychological elements" and identical content would be unjustified (Siegel 2013a). Clearly, then, a great deal is at stake not only in discussing similarity and difference between the beliefs and experiences themselves, but also in discussing similarity and difference between the etiologies of beliefs and the etiologies of experiences.

The rough idea is intuitive: beliefs with certain etiologies have reduced ability to justify, so experiences arising in a similar way should also have a reduced ability to justify. But things get complicated when we ask what counts as an experience's etiology having sufficiently similar psychological elements to the etiology of a bad belief. Unfortunately, we aren't given much to go on, but this amounts to a serious problem, since so much depends on just how this notion is fleshed out.

Siegel alludes to a number of paradigm bad belief formation processes which will be illuminating here. She writes:

"The cognitively penetrated experiences described above arise from processes that mirror routes to paradigmatically ill-formed beliefs: wishful thinking, fearful thinking, jumping to conclusions, and reaching belief through rationally ungrounded associations or negative affect." (Siegel 2013a)

As suggested before, we might think, for example, that Vivek is in the grips of an episode of "wishful seeing". Siegel’s idea is that Vivek is having a visual experience as of a largely pleased crowd that was generated in the same way an ill-founded belief is
generated as a result of a process of wishful thinking, which is a paradigm sort of ill-founded belief.

Note, however, that it's unclear what we ought to consider to be an experience-generating process sufficiently similar to ordinary wishful thinking. The required similarity cannot be too strict or it won't allow processes that result in experiences rather than beliefs, but it also can't be too relaxed. Clearly not just any case in which a desire that $p$ causes a belief that $p$ will do, as this will overgenerate downgrade. Consider Lyons' Good Snakes example. Notably, as Lyons (2011) points out, it doesn't seem to matter to our intuitions what the penetrating state is. GS could query her low-level perceptual representations for accurate snake-template matches more often because she wants to find snakes (perhaps she's a herpetologist desperate for research specimens). As long as the mechanism is the one described in Good Snake rather than the one in Bad Snake, GS still seems justified in believing there are snakes on the basis of her perceptual experience, despite the fact that her desire to see snakes caused her to have a visual experience as of a snake and form a belief with the same content on that basis.

I take the above example to illustrate that the challenge Siegel faces in spelling out the mentalist conditions on epistemic downgrade in cognitive penetration cases is not a small one. She needs to find the right level of description and the right way of specifying the required similarity between processes to successfully capture the bad-making feature, whatever it is, of cases like wishful thinking, and successfully carry that feature into a sorting of experience-generating cases that match our intuitions.

We may note, further, that there is a danger of violating mentalism if the mentalist conditions for epistemic downgrade are not developed further in a mentalist-friendly way.
It could, after all, be the case that what makes wishful thinking a paradigm case of bad belief formation is that it's a route to belief in which a desire that p causes an agent to form a belief that p in a way such that the belief that p is *unlikely to be true*. So, if what counts as a sufficiently similar experience formation process to the process of wishful thinking is that it is a process in which the experience as of p is caused by a desire for p in a way such that the experience as of p is *unlikely to be veridical*, this will help sort the cases correctly, but it will be because it is a covert invoking of reliability considerations.

Further, as Lyons points out, at some levels of description, it is far from obvious that there are appropriate analogs in belief-formation processes to the elements of experience-formation processes (Lyons 2016). In Good Snake and Bad Snake, for example, there is clearly some relevant difference between the cases since we have the intuition that one is good and one isn't, and when we describe the experience generation processes themselves, there are ways of describing them as being different in potentially relevant ways (as they are above). But these differences in perceptual mechanisms might just not have a clear analog in any belief formation process, which Siegel's view requires. Lyons writes:

"The main problem is that there doesn't seem to be any obvious candidate purely doxastic process that's strictly analogous to either the Good Snake or the Bad Snake process. What, for example, is the doxastic analogue of a stored visual template? If the standard canons of rational belief formation are supposed to sort the relevant processes, these won't be processes that involve primal sketches, visual form templates, and the like." (Lyons, 2016)

Siegel's early mentalist account of cognitive penetration cases, it seems, is still underdeveloped, and that development, unfortunately, faces a number of challenges. Unless it manages to overcome these, it seems in a poor position to compete with the
traditional internalist and the externalist. The traditional internalist, on one hand, offers an account that respects an intuition the above mentalist view rejects: that the subject in cognitive penetration cases does seem to have done everything reason demanded of her. The externalist, on the other hand, seems to be in a better position to offer an account which sorts cases in accordance with our downgrade intuitions and explains those intuitions.

**Quasi-Inference: McGrath**

Rather than focusing on the *process* by which penetrated experiences arise and degrees of similarity between that process and the processes involved in paradigm cases of bad belief formation, McGrath's account is focused on content-based analogizing, invoking our intuitions about what belief transitions would count as good and bad inferences were the subject to transition between them.

McGrath asks us to imagine a case of Peter Markie’s in which a gold miner has what Siegel might call a case of "wishful seeing". The gold miner has a visual experience as of a yellowish rock. The gold miner has a powerful desire for gold and really wants to believe it's a gold nugget. Because of this, the gold miner winds up with a visual experience as of a gold nugget (Markie 2006, via McGrath 2013).

The idea is that there was an "input" visual seeming\(^\text{16}\) (the experience as of the yellow rock) and an "output" visual seeming that is much richer in content (the experience

---

\(^{16}\) In this section, I will follow McGrath in using the term “visual seeming”, which I assume to track the same thing as “visual experience”, unless otherwise noted.
as of a gold nugget). The input seeming was "freely enriched" by the penetrator (the wish for a gold nugget) to the richer output seeming in that if the input and output states were beliefs, we would regard the agent as having made a kind of bad inference: she jumped to conclusions, as the belief that there's a yellow rock there and the wish for gold don't support an inference to "there's gold there". Hence, McGrath wants to say that the jump from one visual seeming to the other was a bad "quasi-inference", and as such the output seeming state cannot justify beliefs based solely upon this visual seeming. McGrath says:

"A transition from a seeming that P to a seeming that Q is 'quasi-inferential' just in case the transition that would result from replacing these seemings with corresponding beliefs that P and Q would count as genuine inference by the person." (McGrath 2013)

Some necessary conditions for counting as an inference McGrath favors are:

"(i) the input and output states must be mental states of the person, not merely of a subpersonal system;"

"(ii) there must be an explanation in terms of the person's own mental states which 'rationalizes' the transition, i.e., allows us to see the transition as the person's treating the content of the input state as supporting the output state."

(ibid.)

If the output seeming is the result of a quasi-inference, then the output seeming can justify derivatively if at all – quasi-inference disallows direct justification. Note, again, that factors beyond the agent's ken can still result in a downgrade in doxastic justification on McGrath's account.

McGrath notes that this account seems to get the cases right. Jill's case, for example, might be understood in the following way: First, Jill's visual experience as of Jack's face without any emotion properties represented and Jill's visual experience as of angry Jack-face are both plausibly personal level states of Jill. Second, we can see Jill as treating the emotion-neutral description of Jack's appearance as standing in a relation of support to the
belief that Jack has an angry look. If the states were doxastic states, such a jump between these beliefs seems plausible as an inference: both states are personal-level, and we can tell a kind of rationalizing story. But this would be a bad inference, since the content of the input visual experience and Jill's fear do not together provide enough support for the output visual experience. So, Jill has made a bad quasi-inference, and the resulting visual experience suffers an epistemic downgrade.

What counts as a quasi-inference is admittedly not well-specified yet, and this, of course, needs fleshing out eventually, since the view relies on the distinction heavily. However, there is a more immediate problem. Intuitively, against the first step in the interpretation in the previous paragraph, none of the cognitive penetration cases yet discussed seem to involve distinct conscious perceptual states or a personal-level transition of any kind at all. You might have thought that in the Jack and Jill example, Jill is afraid that Jack is mad at her, she turns her head to look at him, and there it is: an instantaneous experience as of Jack looking angry. After all, if things didn't occur this way, wouldn't Jill notice? Or failing this, shouldn't she notice (that is, might not the phenomenological contribution of the "jump" serve as a defeater)?

One might be tempted to say that these considerations put many of the examples to be explained beyond the scope of this account, and what examples they do explain seem phenomenologically questionable. McGrath, however, says in a brief footnote that the jumping to conclusions between visual seemings (and between beliefs) need not involve a temporal gap, but rather can be at an instant. He says "what matters is that there is the sort of asymmetric dependence associated with inference." (ibid.)
This is an interesting suggestion, but it's unclear whether or not the details of the suggestion can be filled out while preserving the view, as a number of potential problems begin to call out for attention if both the basic and the enriched experiences are occurring instantaneously. To begin with, I take it that the idea is not to attribute to subjects like Jill and the prospector two independent visual experiences occurring side-by-side or occurring as a single complex experience that somehow encompasses two conflicting phenomenologies (one neutral-face phenomenology and one angry-face phenomenology). I take it McGrath must instead want to say that the subject is having an experience as of a gold nugget, and that that very experience is also the experience as of a yellow rock - what we have here is just a single experience. Perhaps the overall experience of a gold nugget has another experience as a proper part (the experience of a yellowish rock): the overall experience might "depend asymmetrically" on the part in this way. Still, more would need to be said about this dependence and why this kind of relation might ground the attribution of an “enrichment” to the experiencer.

Given that, for our theoretical needs, there needs to be this sort of asymmetric dependence relation between an experience as of a gold nugget and an experience as of a yellow rock, one very plausible suggestion for filling in some of the details is to take the dependence relation to be the dependence of high-level properties (e.g., kind properties, or perhaps affordances) on low-level properties (e.g., certain color and surface properties). (See Chapter II for more on “Rich Content”). To see how this might work, consider, for example, the pine tree expert whose visual experience upon looking at pine trees has changed since he gained his expertise. If this is best explained as the expert’s coming to visually experience pine trees as pine trees, there is a relevant sense in which the experience
representing the trees as pines might be dependent upon the experience representing low-level properties. In the expert's case, the low-level properties together with his expertise adequately support his experience of the trees as pine trees, and so the expert isn't guilty of a bad quasi-inference. This isn't the case, however, for the gold miner's experience of a gold nugget. His experience of a gold nugget is dependent upon his experience of a yellow rock, but the support-gap between the yellow-rock-seeming and the gold-seeming was just a leap fueled by desire - the former doesn't support the latter.

However, if a picture like this is what McGrath has in mind, it seems to me that there are still a couple of worries looming.

First, in Chapter II, I argued against "Rich Content" views according to which properties like kind properties are represented in conscious visual experience. Even putting this aside, it is at least clear that rich content is rather controversial, and a theory should be less appealing insofar as it requires contentious premises to get off the ground.

Further, the intuition that something is epistemically amiss in cases of cognitive penetration isn't sensitive to the presence or absence of dependency relations among levels of experiential content, making an account couched in terms of those relations seem unlikely to be able to completely explain the intuition. Even if a strictly low-level content view is right (we only represent colors, facing surfaces, etc. in experience), there still seems to be room for cognitive penetration examples, and these examples don't seem to me to be intuitively different in kind from ones that include facial-emotion and gold-ness as represented in perceptual experience. Imagine I irrationally believe that bananas are blue, and when I see a banana-cutout, my experience is penetrated by my false belief the bananas are blue, resulting in an experience as of a blue banana-cutout whether or not the cutout is
actually blue. This seems to me, intuitively, to be of the same kind of case as at the top. However, it's unclear what the input and output visual seemings would be in this case.

In fact, research purporting to show just such low-level penetration effects exist show up in the empirical literature, and these are the ones that seem to be the most difficult to explain away as post-perceptual. Levin and Banaji, for example, found that participants tended to rate a grayscale image of a white male as “lighter” in color than a greyscale image of a black male, even though the average luminance of both images was set equal (Levin and Banaji 2006 - For other examples, see Delk and Fillenbaum (1965), Hansen et al (2006) for a purportedly real banana-color example, see also MacPherson (2012) for an overview of some such research). One might think that if cognitive penetration happens at all, the strongest cases for it are to be made at the low level.

Mentalism Part 2: Siegel (2017)

Last, let us turn to Siegel's most recent view. Siegel's Rationality of Perception is an expanded project which represents the development of Siegel's earlier work in the area. In addition to an expanded scope and a number of explorations into nearby areas, the new work is different in two major ways that are relevant to the present discussion, and I will examine each in turn here, ultimately concluding that the most recent version of Siegel's mentalist view still faces significant difficulties. The first major difference I'll discuss is that, in Siegel's most recent view, perceptual experiences don't just have different justifying powers dependent on their etiology. Rather, Siegel now argues that the experiences themselves can be rational/irrational, and they redound on the rational
standing of the experiencer. The second major difference: in Siegel's earlier view, she argued that we ought to accept that there is epistemic downgrade in cases like Vivek's and Jill's because of a high degree of "similarity" between the etiologies of experiences like Vivek’s and Jill’s and the etiologies in several paradigm bad belief formation cases. In her most recent view, however, she moves away from couching the account in terms of similarity, and argues that we should accept that there is downgrade in cases like Vivek's and Jill's because the perceptual experiences in those cases are the result of inference, and that we should, accordingly, afford perceptual experiences rational standing of their own, as anything that is the result of inference must be rationally evaluable.

*The Rationality of Perception*

The first new feature to discuss is the claim that perceptual experiences can *themselves* be rational/irrational. According to Siegel, experiences themselves are rationally evaluable, can arise through rational/irrational processes, and redound on the rational standing of the experiencer. On this view, when Vivek looks out onto the lukewarm crowd and has a vanity-penetrated visual experience as of smiling faces, not only is Vivek irrational in forming or strengthening beliefs on the sole basis of that visual experience, but the visual experience *itself* is irrational, Vivek is irrational in having it, and all this because it was generated by an irrational process.

As the title of her monograph communicates, this claim is of central importance to Siegel’s most recent view. She calls the claim "The Rationality of Perception" (and I will follow her in this here). She puts it this way:
"The Rationality of Perception: Both perceptual experiences and the processes by which they arise can be rational or irrational." (Siegel 2017, p.15)

The Rationality of Perception is primarily presented as a solution to the problem of deciding how we ought to think about the epistemic situation of those like Vivek and Jill, similar to the problem being addressed in Siegel’s earlier view (Siegel 2013a). Siegel's most recent view is, like her older view, a mentalist one on which the perceptual experiences of Vivek and Jill have downgraded justifying ability. On her most recent view, however, what explains this is that the experiences themselves are irrational, and the beliefs solely based on these irrational states are irrational as well, in the familiar way that beliefs based solely on irrational beliefs are themselves ill-formed. She writes:

"[I]t is not rational for Jill, Vivek, or the preformationist to believe their eyes. It is not rational, because their hijacked experiences are irrational. They are irrational because they are epistemically sensitive to their psychological precursors, in the same way that conclusions of inference epistemically depend on inferential inputs." (Siegel 2017, xxiii)

Just how does this change the dialectical situation? That is, what does the mentalist have to gain by trying to show that perceptual experiences themselves can be rational/irrational, rather than holding that although they are not rationality-apt, they can still have different amounts of justifying power based on their histories?

Siegel mentions this difference, but mostly avoids specifics, focusing instead on broader theoretical reasons for which inquiry into the Rationality of Perception thesis might wind up illuminating, writing:
"Both positions, I found, have a lot of explaining to do. And the more I explored considerations against the idea that experiences can manifest an epistemic status like justification, the less powerful those considerations seemed." (ibid, p.xxiii)

Siegel does offer, in that same passage, that "[t]he less measured, more extreme view offers a cleaner account of the epistemic situation than the moderate view" (ibid, p.xxiii), but doesn't elaborate there on the respects in which the view comes out cleaner.

One advantage, however, that has been brought out in the present work, is that Siegel avoids one of my earlier standing criticisms. Recall that the view presented in Siegel’s earlier work was largely defended there by pressing an analogy between etiologies of belief and etiologies of experience. Since certain etiologies of belief can result in beliefs with reduced justifying power, Siegel argued, experiences with "similar" etiologies should result in experiences with reduced justifying power, as long as no relevant disanalogy can be found between beliefs and experiences. Siegel argued against the relevance of a number of disanalogies between experience and belief which she took to be the most plausible differences a critic might cite as reason to treat experience and belief differently with respect to justifying power (Siegel 2013a). However, in Section 4 of the present chapter, I argued that, in particular, her defense in one case was insufficient. It is, I argued, still open to a critic to hold that the relevant disanalogy that undermines (or is at least a good place to start investigating what undermines) the claim that we ought to consider the etiologies of both beliefs and experiences in determining justifying power is the fact that beliefs can themselves be rational/irrational while experiences cannot. On the present view, this problem (obviously) disappears.

The case Siegel makes for accepting the Rationality of Perception, however, faces a number of its own difficulties. Siegel examines what she takes to be the most powerful
reasons for thinking that experiences are precluded from having a rational standing. She argues that all of these fail to threaten the Rationality of Perception claim, concluding that we lack good reason to deny rational standing to experiences. I will discuss her responses to several potential objections here\(^\text{17}\). Note that because she is arguing in this way, Siegel cannot afford for a single defense to fail. In her first two cases, I simply present the potential reason to deny rational standing to experience and Siegel's response in order to show my own suggested reason to deny rational standing avoids her objections. In discussing a third reason to deny rational standing, Siegel's response splits into two parts, and I criticize both. I then proceed to suggest another potentially relevant difference between beliefs and perceptual experiences, which, I argue, makes the right assignments of rational status in the cases raised by Siegel.

Siegel notes that the most powerful thread of potential objections to the Rationality of Perception come from the idea that perception is in some sense "passive" quoting Ernest Sosa as an example of one such objector (see Sosa 2007). We already saw that McGrath's account attempted to respect something like this in building "passive" and "active" categories into his view, discussed earlier, in the form of experiences that result from quasi-inference, which justify derivatively, and those that don't, which justify directly. Siegel proceeds to propose a number of different kinds of "passivity" one might take to be threatening to the Rationality of Perception and rejects each one, thus securing the Rationality of Perception from the "passivity" worry.

\(^{17}\) I omit "passivity with respect to reasoning", as it seems to me this is covered in discussion of "passivity with respect to reflection", and I discussion of the way in which experience is "not adjustable" by "habituation", since Siegel's responses is to argue that it is adjustable in that way (and I agree, since knowledge of visual learning might enable an agent to do this).
First, she considers and rejects a phenomenological passivity that amounts to the fact that "[i]t is not part of the phenomenology of perception that our experiences seem to result from mental activity of any sort." (Siegel 2017) This, she argues, can't preclude experiences from having rational status, as there are many beliefs which clearly hold rational status but which lack a “mental activity result” phenomenology. This seems right. She writes,

"We simply find ourselves believing that it is time for lunch, that the audience is pleased, that our neighbors are kind, or that the music is too loud." (Siegel 2017, p.32)

Next, she rejects another kind of "passivity" which she calls passivity with respect to "reflection". Reflection, for Siegel, refers to "upper-echelon rational capacities". These capacities include “the capacity to deliberate about what to believe, to formulate one’s reasons for what one believes or decides, and the capacity to revise or adjust one's conclusions in light of reflection on other things one believes or one's reasons for holding the belief." (ibid. p.32) Siegel suggests that the opponent of the Rationality of Perception might argue that every belief could, in principle, be formed by reflection and so beliefs are active in this way, while this is not the case with perceptual experiences, and so experiences are passive in this way.

Against this, Siegel argues that a toddler who has just put on her socks with great effort is a case in which a subject believes her socks are on through a non-deliberative route (and for whom deliberative routes to belief are not a live option). She argues that, nonetheless, have the intuition that the toddler's belief about her socks deserves rational standing. She writes:
"The route to belief wasn't deliberation, which would require more self-conscious reasoning than she is capable of, but rather a mix of observation integrated with action. A more mature subject could reach the same first-person belief ... by reflection. The toddler couldn't, yet her belief seems to contribute to her rationality just as much as the older subjects' beliefs contribute to theirs." (ibid, p.33)

Admittedly, it is somewhat unclear to me what Siegel counts as "reasoning" in this case, but whatever kind it is, we can at least take away the following: it does seem as though any proposed "passivity" that would grant rational standing to the sock-beliefs of the adult, but not to those of the toddler has a strike against it. So, for the purposes of this discussion, we can let "reasoning" in this case refer to any account of reasoning which would do this and attempt to avoid this strike.

Last, Siegel discusses the claim that experiences cannot have rational standing because they cannot be rationally adjusted. She writes:

"According to this idea, for a mental state to be epistemically charged, it must be possible for a subject to adjust it, in order to make it conform to any epistemic norms that can be used to evaluate it. It doesn't seem possible to adjust your experience, other than by looking away, covering your eyes, or otherwise closing off perceptual input" (ibid, p.34)

Siegel considers two main aspects of what might be called "adjustability", and concludes that neither could be grounds for denying rational standing to perceptual experience.

These are of particular concern to us here, and I will criticize Siegel's defenses against both. First Siegel examines adjustability by "disowning". She, then, discusses adjustability by "deliberating".

First, she considers what she calls adjustment by "disowning a mental state". She argues that experiences and beliefs are on a par in this respect, since one can cease to rely
on experience in reasoning and action in the same way one can cease to rely on a belief in reasoning and action. She, however, also admits that there is a difference in this area, writing:

"In the case of belief, ceasing to rely on a belief can't come apart from ceasing to have the belief. But experience can persist, even if you don't use it in reasoning or action. If the experience persists, does that show that experiences are never fully rationally adjustable?" (ibid. p.35-36)

She argues that the answer to this is "no", writing:

"Consider the Müller-Lyer illusion. If you know the lines are not the way they appear, and you cease to rely on the experience, then there is no further rational adjustment to be made in the situation... " (ibid, p.36)

This response is somewhat puzzling. It seems as though when we're talking about which differences between experience and belief might be relevant to whether or not experiences are rationally evaluable (and whether one can be rational or irrational in having them), it's quite plausible that it's relevant that ceasing to rely on a belief in reasoning and action counts as extinguishing the belief, but one still counts as having an experience no matter what one does with it with respect to reasoning and action. After all, this is a difference in the conditions under which one counts as having the target state. Whether or not there is "further rational adjustment to be made" seems to depend on what we mean by "to be made". If this means that no further rational adjustment is available to the subject, then this seems right, but then this doesn't address the original question: in the case of belief there are no remaining adjustments available to the subject and the subject, thus, no longer has the belief, while in the case of experience there are no remaining adjustments available to the subject, but this is unrelated to whether or not the
subject is having a perceptual experience. This still seems like a potentially relevant
difference. If, on the other hand, Siegel means that the viewer is under no rational
pressure to adjust anything other than ceasing to rely on the Müller-Lyer experience in
reasoning and action, this seems friendly to her opponent, who might then suggest that
the subject faces no rational requirements involving the Müller-Lyer experience because
the subject's experiences have no bearing on the subject's rational standing.

Siegel elaborates with an example:

"Contrast Vivek's experience when it arises from his vanity. If Vivek learned that
his experience arose from his vanity, and the experience persisted because of his
vanity, that would be a case of residual irrationality. Vivek's situation would be
analogous to someone obtuse who disowns an attitude (e.g. disrespect for
someone they treat badly), but lacks the understanding needed to correct all the
perspectives that go with it. Sometimes the fact that an experience persists, even
when a subject disowns it, constitutes this kind of residual irrationality." (ibid.
p.36)

This response is somewhat difficult to interpret. Vivek, I presume, is mentioned here to
stand as an example best understood according to the view Siegel is attempting to argue
for. Vivek is someone disowning a perceptual experience who, according to Siegel,
should count as irrational in having that experience despite the disowning. He learns his
experience is caused by and is sustained by his vanity, and now means to quarantine the
experience, cutting it off from reason and action, just like the viewer in the Müller-Lyer
example, but in having the experience nonetheless, Vivek is (residually) irrational, unlike
the Müller-Lyer viewer.

So, at least as I reconstruct it, Siegel’s argument is that the fact that one cannot
cease to have experiences the way one can cease to have beliefs only seems to indicate
the experiences might be precluded from having rational status when we examine
innocuous cases like the Müller-Lyer case. We're thinking "there can't be anything wrong with continuing to have that experience" because, even on Siegel's account, there isn't anything wrong with it in this case, but this is fooling us into thinking there can't be anything wrong with having a perceptual experience generally. However, in cases like Vivek's where the experience is the result of Vivek's vanity, we ought to hold that the experience of the happy crowd is residually irrational for Vivek to enjoy, even though he's behaving appropriately at the cognitive level.

This argument, however, would have to be made on the strength of the comparison between Vivek's case and the case of the person disowning an attitude like disrespect (mentioned in the quote above), since the idea is to argue that granting rational status to perceptual experiences is an independently attractive view, on the strength of which one might argue that cases of cognitive penetration should result in ill-formed beliefs. It is, however, difficult to see how the disrespect-disowning example provides much support in this case. The “attitude of disrespect”, I take it, is supposed to be a case in which there is a state which we uncontroversially regard as irrational, but which one does not cease to have even when one has ceased to rely on it in reasoning and action. In describing this case, however, it’s an “attitude of disrespect” that is being disowned, but it's "perspectives that go along with it" that remain nonetheless. If these perspectives count as constituents of the attitude of disrespect, then the subject hasn't really disowned the attitude (perhaps, he intends to, but fails), and so the case can't show us anything about cases where a state is "disowned" but still persists. If the perspectives that remain do not count as constituents of the attitude of disrespect, then the most natural way of filling out the story seems to be that the attitude of disrespect is actually gone, and that
what remains are merely behavior patterns and other attitudes that developed in response to the attitude of disrespect. The subject may, of course, be irrational in virtue of some of those behaviors or attitudes persisting, but since the subject no longer has the “attitude of disrespect”, she can’t be irrational in virtue of that attitude’s persisting. Again, this is what would be needed to argue that one can be rational/irrational in having a state which is “disowned” but still, itself, persists. I conclude that it remains open that the way in which one does not cease to have experiences when one ceases to rely on them in reasoning in action might mark a relevant difference from belief that precludes experience from harboring rational standing.

The other kind of adjustability Siegel examines is adjustability by deliberating about what to believe. Siegel accepts that experiences do not have this feature, writing:

"If experiences were adjustable by deliberation in response to rational criticism, then in response to criticism, such as the information that the contents of one's experience is heavily influenced by vanity, or that the experience is irrational, one would have to be able to explicitly reason to a new experiential conclusion that rationally addresses the critical information." (ibid, p.34)

Siegel argues that lacking this kind of adjustability should not preclude experiences from having rational standing, as there are beliefs we think have rational standing but which also lack this kind of adjustability.

Her main example is a sufferer of Capgras syndrome who firmly holds the delusional belief that some of her friends and loved ones have been replaced by excellent impostors. She writes:

"Beliefs in delusions cannot be adjusted by deliberation. Rarely if ever has anyone been talked out of the monothematic delusional belief in Capgras syndrome, for example, or out of delusional beliefs in schizophrenia. The background pathology
prevents these beliefs from being revised. These beliefs seem to be paradigms of irrationality." (ibid, p.35)

What to think about this example, however, as well as other nearby cases such as cases of schizophrenia, depends on how the details of the story are filled out. For example, there's some good reason to think that actual Capgras patients suffer a disruption in the connection between ventral stream areas of visual processing and the limbic system, making it plausible that the problem is, at least in large part, some sort of failure of emotional arousal or feeling of familiarity in response to visual cues. When combined with the failure of a "global consistency checking" mechanism and associated problems linking portions of episodic memory together, this goes a good distance toward explaining the behavior seen in Capgras patients (Hirstein and Ramachandran 1997). But this doesn't seem analogous to the case of perceptual experiences, where no amount of reasoning could change the experience's content. Rather, this seems to be a scenario in which the evidence available to the subject is very strangely gerrymandered. It seems plausible that, for the subject, the strange new patterns of feelings she gets in the presence of loved ones are taken as evidence that the loved one is an impostor. Perhaps, in addition, failure to register connections in episodic memory deprives the subject of a source of evidence that the person in front of her is, in fact, the same person who taught her to tie her shoes, and ride a bike, etc. If the feelings had been different, or if she had been able to link episodes in memory, she might not have thought the person in front of her was a father-impostor. Of course, that may not be all there is to the story. Perhaps there's a sense in which the subject is placing too much weight on the lack of feeling she gets while looking upon her father. The main thrust is that it seems possible in cases like
this that the delusional beliefs might be very robust, just as Siegel cites, but without also
being, in principle, inaccessible to critical reflection the way visual experiences are. In
fact, anecdotes associated with subject DS seem friendly to this interpretation.
Fascinatingly, it seems that DS's father was able to ease DS's Capgras symptoms
somewhat for short periods of time by telling him "the man who you have been with all
these days is an impostor - he isn't really your father. I have sent him away to China. I am
your real father - it's so good to see you son." (Ramachandran and Hilstein 1997)

Further, we ought to note that it's quite plausible that our intuitions about the
rational assessability of Capgras beliefs is counterfactually dependent on the
malfunctioning cognitive systems, some of which are mentioned above. That is, we might
think that the beliefs are adjustable in the sense that if the subject's cognitive faculties
(like episodic memory and consistency checking) were functioning properly, rational
deliberation would change them. Clearly, however, there is no analogous move one could
make in the case of ordinary functioning perceptual experience.

Siegel, it seems, anticipates this response, suggesting that in Capgras there is a
"background pathology" one might think is separable from the beliefs themselves, but "in
schizophrenia, it makes little sense to distinguish the beliefs from the background
pathology that precludes their adjustment." (Siegel 2017, p.40)

I have a rather poor sense of what should and should not count as a case of
schizophrenia, and so, accordingly, I do not have strong intuitions about whether or not
cases of schizophrenic belief should count as cases of irrational belief, insofar as
suffering a certain kind of schizophrenia just is having a certain belief. But, of course, we
needn't restrict ourselves to likely actual cases. So the general consideration I wish to
advance is the following: it seems to me that whenever we attempt to fill out the details of a hypothetical delusional belief, the more it's the case that the subject just finds herself holding a belief which is completely inaccessible to adjustment processes, the more comfortable it becomes to think of it as one of the non-rational "zap" cases described in Siegel’s earlier view (Siegel 2013a) - cases where you just get struck by lightning or have a terrible flu and the sudden result is a set of non-rationally acquired beliefs. My claim is that in the cases of delusion that resemble actual cases, the delusional belief seems rationally assessable only because it seems plausible that the belief actually is sensitive to deliberation, at least in principle. On the other hand, it is very difficult, I think, to imagine a belief that is really, truly acquired and maintained with complete insulation from any of the agent’s deliberation or evidence. It’s not quite clear to me that such a case would still count as the agent holding that belief – details would be needed here. But insofar as such a case is coherent, it seems to me the temptation to classify the belief as rationally assessable diminishes sharply once we eliminate all connection between the other states of the agent and this completely isolated belief.

It may be the case that, in the discussion above, I've focused on too broad a notion of adjustability. It sometimes seems that what Siegel had in mind by "adjustment by deliberation" involved only very specific, conscious, high level deliberation processes. In that case, it seems like Siegel's toddler cases suffice to show that this kind of lack of adjustability shouldn't prevent experiences from having rational standing. We shouldn't deny that a toddler's beliefs have rational standing, but they certainly can't adjust their beliefs by deliberation since they can't deliberate, in this high level, self-aware way, at all.

---

18 In Siegel’s earlier view, “zap” beliefs were taken to be non-rationally acquired, justified beliefs (Seigel 2013a)
But then, it also seems like "adjustable" might be broadened to include the kind of adjustment toddlers do make, and the argument still seems to stand. Toddler beliefs are in some sense adjustable – Siegel’s story about the toddler is an excellent example of this.

While the toddler, perhaps, does not consciously deliberate, the toddler’s beliefs definitely change as a response to a variety of the Toddler’s incoming states, and they do so in a way that allows us see the toddler as responding to the incoming states as a kind of rational support for those beliefs. Though the toddler cannot perform conscious deliberation, it still seems quite plausible that the toddler is, in some sense, taking the experience of putting her shoes as evidence for the belief that her shoes are on, and might have formed a different belief had she had different experiential evidence. The toddler’s visual experience as of her socked feet, on the other hand, won’t change no matter how much experiential sock-putting-on evidence she does or doesn’t have. It’s not clear to me what we ought to call the process by which a toddler forms beliefs on the basis of her experiences, etc., but it still seems clear that there are problematic differences adjustability differences between a toddler’s experience and a toddler’s ordinary beliefs even if we abstract away from a very high-level, self-aware account of “deliberation”.

Having discussed Siegel’s responses to several proposed experience/belief differences, I now wish to suggest an additional experience/belief difference that might preclude experience from manifesting rational standing. All the differences mentioned are examined case-by-case, the idea being that we're examining a number of features of experiences/beliefs and determining whether or not each one is such that any state lacking/possessing that feature cannot be rational/irrational or redound on the rational standing of its subject. However, it also seems plausible that what kind of states we allow
into the "house of reason" might be determined by properties about that state type or the mechanisms that produce that state type. For example, it might be the case that we think experiences are precluded from having rational standing because, in general, experiences are very robustly insensitive to the states of evidence the subject possesses. Beliefs, on the other hand, are generally highly sensitive to the states of evidence the subject possesses. If beliefs were generally deeply insensitive to the subject's evidence (as long as there is sense to be made of this notion), it seems plausible that they might not be rationally assessable, and if experiences were as generally deeply sensitive to a subject's evidence, they might get rational standing.

Note that this gets the cases proposed by Siegel, above, right. Sensitivity to evidence does not invoke conscious deliberation, and so does not disallow beliefs formed without an effect on one's phenomenology. It does not invoke deliberation, and so does not preclude toddlers' beliefs from redounding on their rational standing. The way in which experiences cannot be extinguished, which we found to be a potentially relevant difference, is a way in which experiences lack a certain kind of sensitivity to states of the subject that beliefs possess. Last, "adjustability", which we also found potentially relevant, is quite close to the same claim I am making here (my suggestion abstracts away from the abilities of the agent).

Note also that this suggestion makes me hostage to empirical fortune, in a certain sense. If it turns out conscious perceptual experience is well-encapsulated, then experiences won't count as having rational standing. If, however, cognitive penetration turns out to be pervasive, then experiences will count.
I think this is not an inappropriate feature for a view in this area to have. One of the main reasons for the philosophical interest in this issue is that might turn out that we've been giving perceptual experience an epistemic free pass it didn't deserve all this time. It’s not implausible that this could be something we discover with the aid of empirical research, especially if that research revealed that perceptual experience works in a way that diverges sharply from our naïve theory. So, I happily accept that my suggested criterion for rational assessability will give different verdicts on the rational assessability of perceptual experience depending on how sensitive experience turns out to be to certain kinds of states of an agent. Siegel, it seems to me, has a similar commitment. Since she ultimately argues that states are rationally assessable if they’re the product of inference, Siegel’s account will similarly give different verdicts in a way at least partly dependent on the empirical literature once a well-defined notion of “inference” is in hand.

With respect to my own case, it happens that there is great wealth of published work aimed at showing cognitive penetration occurs in a great variety of forms, but this doesn't mean I'm proposing we accept the Rationality of Perception on the basis of perceptual experience's sensitivity to the subject's evidence.

First, I've already argued against "rich content": I argued against high level properties’ (like kinds, gender, the happiness of an audience, or the anger of a friend) figuring into the content of conscious perceptual experience. If this is correct, then this reduces the potential for cognitive penetration quite a bit. If rich content were a part of perceptual experience, cognitive penetration might be everywhere: you might look at a picture and have an experience as of a stylish man-shape, and then learn that you’re looking at a picture of Elvis Presley, and then suddenly you’re having a visual experience
as of Elvis because you were just told who the person in the picture was. Similarly, a portion of the empirical literature is aimed at showing that cognitive penetration occurs at this high level, but if there is no rich content in perceptual experience, then this portion of the literature turns out (likely) to be investigating post-perceptual effects. Having argued against rich content, the potential cases of cognitive penetration are restricted to cases in which we have reason to believe that some cognitive states of ours penetrate into conscious visual experience and change the low-level features of that experience like color and shape.

Further, as Scholl and Firestone argue, despite the large number of recent publications aimed at producing cases of cognitive penetration in controlled settings, whether or not cognitive penetration is really occurring in those settings is still a wide-open question, thanks to a small number of pervasive confounds like low-level display features, attentional effects, and task demands shared even by some of the most striking and oft-cited work (Firestone and Scholl 2016).

For example, the Levin and Banaji (2006) article (cited in both (Siegel 2017) and in the present work) is one of the most striking demonstrations of a proposed cognitive penetration effect, wherein a pair of images of human faces are presented with their average overall luminance and contrast set equal. Participants, however, generally rated the face with stereotypically white features as lighter-looking than the face with stereotypically black features. However, Scholl and Firestone found that blurring the images enough to cause subjects to seem not to categorize the images racially (subjects reported answered a series of questions during which they reported things like believing the two images to be images of the same person, etc), did not remove the darkness
judgment effect. This, they claim, suggests that the effects are resulting from low level features of the display (such as cues to lightness that are independent of average luminance/contrast, such as the depiction of patterns of illumination and shadow), rather than perceived racial categories. (Firestone and Scholl 2015)\(^{19}\)

I conclude that it is still quite open whether perceptual experience is such that the experiences are not themselves rational or irrational, and that they do not redound on the rational standing of the subject.

**Inference**

The other new feature of Siegel's view is an account of inference on the basis of which Siegel argues we have positive reason for thinking of perceptual experiences as rationally evaluable. If experiences can be the result of inference then they must, like beliefs, be capable of manifesting rational status. She writes:

"When one arrives at a conclusion by inference, the conclusion's epistemic power to support subsequent beliefs, as well as its own epistemic standing, can be modulated by the inputs to the inference. Inference is an example - perhaps it is the prime example - of a process that can issue in beliefs that can be rational or irrational. If perceptual experiences can arise from inference, then the Rationality of Perception is true." (Siegel 2017, p.19-20)

Of course, the next issue is whether or not perceptual experiences *do* count as issuing from inference in the correct way. What counts as an inference? Siegel begins with what she calls the "Response Hypothesis":

---

\(^{19}\) See (Firestone and Scholl 2016) for many, many more.
"[I]nferring is a distinctive kind of response to an informational state, or to a combination of such states, that produces a conclusion. An informational state can take the form of a belief, but it can also take the form of a perceptual experience, a supposition, or an attitude (in the psychologist's sense)." (ibid, p.77)

Siegel aims only to provide "illumination" with respect to the particular sort of response inference is, and not "analysis", and admits that this illumination will not help us determine in hard cases whether or not a particular response is an inference (ibid). So, of course, the explication of "inference" is less complete here than ideal for the theoretical goals being pursued. However, the rough picture that emerges is of a certain, special kind of response to a state of information that need have neither a phenomenological profile, nor a "reckoning state" which counts as the subject recognizing the input supporting the conclusion in some way. The picture Siegel gestures toward is one friendly to counting at least some experiences as the output of inference.

Once again, Siegel defends experience as the result of inference by suggesting potential disanalogies between certain responses resulting in beliefs and certain responses resulting in experiences, and arguing that none provides good reason to think them not on an epistemic par. The one I wish to examine here is the claim that a response's occurring at a sub-personal level does not detract from its status as an inference.

Siegel considers two paradigm sub-personal process types: early vision processes, and grammatical processing. She concludes that what they have in common is that they are widely thought to be insulated from information in the rest of the mind, and in this sense form a subsystem. Siegel, however, focuses on showing that an input's being informationally isolated doesn't detract from the associated response's being an inference. She writes:
"There is ... evidence that some of our most basic assumptions about the physical world and social world are informationally insulated - such as our assumptions about how objects move through space (they move as bounded wholes), which spatio-temporal processes are processes of mechanical causation (the ones without big spatial or temporal gaps), and which events are exercises of agency. It is far from obvious that these working assumptions of our cognitive and perceptual systems are lacking in epistemic power or that they are exempt from epistemic appraisal." (ibid, p.104-105)

As far as showing that an insulated state of information might still be the input to an inference, this seems right to me. My concern, however, is not that certain isolated inputs (like compartmentalized beliefs) cannot have, as Siegel puts it, "epistemic power" and so cannot be the input to a process of inference. Perhaps it is the case that the stored generalizations of our perceptual system should count as inputs to an inference that ultimately results in perceptual experience for the subject. Rather, my concern is that if transitions like these count as inferences, they may be inferences that have nothing to do with the rational standing of the subject herself. Rather, they seem to be inferences of the subsystem. Experience arises from a process the inputs to which seem plausibly to count as information, and one's visual system does seem to be responding to this information in Siegel's sense of "respond". However, these states of information aren't states of the system as a whole - they're not states of the subject. As Siegel argues, it's true that they're truth-apt informational states and no feature of the states themselves precludes them from entering into inferences in a general sense, but as it happens, they're states of the subsystem rather than states of the subject. I mean to point out not merely that the states of the early visual system are states of the subject that are very far from the reaches of her awareness (that is, these states could not easily enter the subject’s awareness the way many compartmentalized beliefs can). Rather, I mean to point out that they are not even
potentially states of evidence the subject possesses, just like the information encoded by my retinal neurons is not information I possess.

My suggestion here, against Siegel, is that even if certain informational transitions do count as inferences, this would not be sufficient for the outputs of those transitions' redounding on the rational standing of an associated subject. The subject, instead, might merely have a subsystem which is performing better or worse inferences that do not redound on the rational standing of that subject.

Reliabilism

Recall that, earlier in this chapter, I argued that reliabilism is on better footing than the "similarity"-based view presented in (Siegel 2013a) in accounting for our intuitions that there is epistemic badness in cases like Vivek's and Jill's. Siegel takes the inference-based account she presents in (Siegel 2017) as a favorable alternative to her earlier similarity-based account. She writes:

"In Siegel (2013a), ... I suggested that experiences would be epistemically downgraded if their relationship to their psychological precursors was sufficiently similar to a relationship between those same precursors and belief, where that latter relationship would make the belief ill-founded... The inferential approach provides a way to develop this idea that avoids the need to appeal to similarity at all, by directly identifying inference as the relationship between experience and a precursor that modulates the experience's epistemic status."

This new mechanism of inference leads her to propose the following argument against reliabilism's ability to explain our epistemic downgrade intuition. She writes:
"Reliabilism seems to lack the resources to identify the epistemic problem with badly circular inferences that lead to epistemic downgrade. Consider first an uncontroversial case of inference from beliefs to other beliefs. Suppose S infers belief Y from belief X. Then she strengthens her belief X, in response to her belief Y ... But now suppose that before belief X was strengthened in response to Y, it was formed by a highly reliable process. The process of strengthening X in response to Y does not seem any less truth-conducive than the process that generated X in the first place. It seems equally truth conducive: whatever process made belief X true across the relevant range of circumstances before it was strengthened would seem to make it true across the same range after it is strengthened. Yet it is wrong that the original belief X and the strengthened belief have the same epistemic status." (ibid, p.126)

The idea here is that by identifying a certain kind of bad inference (circular reasoning) Siegel can honor the intuition of badness when one infers Y from X and then strengthens one's belief in X on the basis of Y. X was initially formed by a good inference. X was strengthened as a result of a bad inference. Process reliabilism, however, seems to Siegel unable to account for the situation in a similar way. If the process that formed X was highly reliable, then X is just as likely to be true when the strengthening occurs. On a certain understanding of process reliabilism, this would mean that the process of strengthening X in response to Y would have to have yielded a belief in X that was at least as well justified as the original belief in X, since it produces beliefs that are at least as likely to be true. This over-attribution of justification results from the fact that process reliabilism type views aren't typically couched in terms of degrees of belief, and so having justification is a matter of maximizing the likelihood of producing true beliefs, rather than a matter of producing a proportion of true beliefs that matches the degree to which they’re believed.

In Epistemology and Cognition, Goldman argues against the psychological plausibility of degrees of belief. However, elsewhere in the same work, Goldman
suggests (and then rejects) a view he called the "calibration view", meant to handle degrees of belief. On this view

"[A] credence-forming-process is "well-calibrated" just in case for any \( n \), approximately \( 100n\% \) of the propositions in which the process causes us to have a credence of \( n \) in are true" (Goldman 1986)

So, for example, if my belief-forming-process has only ever been activated four times, and each time resulted in a credence of 0.25 for each member of \{p, q, r, s\}, then it's perfectly calibrated if only one of them (25% of the set) was true.

The view, of course, has its own share of potential problems, but these are beyond the scope of the present work\(^{20}\). I mean only to show that such an account is a way for reliabilism to have the kind of sensitivity required to avoid miscategorizing cases like the bad circularity case above.

Let us start with the initial belief, \( X \), which has been formed by a reliable process. Let's say our resulting credence in \( X \) is 0.7. By hypothesis, this belief is justified, so on the calibration account, this initial belief is justified in that the process that produced it assigned/would-assign a credence of 0.7 to a set of propositions that's 70% true (which includes \( X \) as a member).

Next, we infer \( Y \) from \( X \), and then strengthen \( X \) on the basis of \( Y \). Let's say that this results in a credence of 0.9 in \( X \). Recall that the problem was that "the process of strengthening \( X \) in response to \( Y \) does not seem any less truth-conducive than the process that generated \( X \) in the first place. It seems equally truth-conducive" (Siegel 2017, p.126).

\(^{20}\) It seems, however, that at least the reason for which Goldman rejected the view after initially proposing it might be met by allowing the reliability of a process to be assessed not based only on actual cases in which it operates/operated, but how it would operate in nearby possible worlds as well. See (Dunn 2015).
With a "calibration" account of reliable processes in hand, however, it seems like we can say at least that there's room for the appropriate kind of downgrade. That is, the circular-process is likely to be less well-calibrated than the, by hypothesis, very well-calibrated original method of producing X in the first place. Since the circular belief process referred to in this case is just a process where a subject strengthens any belief X produces regardless of what it is, and X is by hypothesis produced in a very well-calibrated way (let's say perfectly), the circular process is bound to assign credences that are too high to match the sets they're assigned to. As per our earlier example, it will assign a lot of 0.9 credences to sets of propositions that are about 70% true.

I conclude, then, that reliabilism is still just as much a live option in accounting for our epistemic downgrade intuitions in cases of pernicious cognitive penetration.

**Conclusion**

I have argued in this chapter that cognitive penetration cases rouse seemingly conflicting epistemic intuitions. I considered a traditional internalist take on cognitive penetration cases as well as a traditional externalist take, concluding that while each offers resources with which one might explain one of the intuitions, neither account contains anything that ought to extinguish the force of the seemingly unfriendly intuition. I, then, examined a variety of recent "mentalist" accounts, according to which epistemic evaluation is still a matter of what's in the head, but cognitive penetration cases like Jill's come out epistemically bad. I argued that each of these accounts, for different reasons, fails to be a more attractive option than reliabilism in accounting for the intuition that something is
wrong with beliefs formed like Jill's, and that all three similarly do not succeed in 
extinguishing or explaining away the force of the traditional internalist sort of intuition 
cognitive penetration cases arouse. I conclude that it remains most desirable and most 
promising to develop a theory that can account for both intuitions.
Closing Remarks

Recall that the first main motivation of this work was metaphysical and empirical.

Chapter I and Chapter II revolved around the question: “What is the nature of perceptual content?” Is perceptual representation fundamentally a matter of representing affordance properties? Does this threaten the idea of perceptual representation generally? What kinds of properties are represented in conscious visual perception? How “rich” can the content of conscious visual experience be?

In those two chapters, I’ve taken steps toward answering these questions. I argued that ecological psychology can be understood as showing which properties in the world are the ones with which we make perceptual contact, and does so in a way that undermines the claim that perception is fundamentally of affordance properties. This result can be used to resist the claim that perceivers are perceptually “locked in”, representing only organism-relative properties. I also argued that the “rich” versus “thin” content debate must proceed in step with the empirical research. I argued that empirical research will generally identify sets of low-level properties coextensive with any high-level properties it investigates, thus delivering a plausible thin content explanation of any high-level contrast or effect elicited in the research. (We will have sets of low-level properties we know the subjects are tracking and which guide their high-level property
attribution behavior.) I argued that in each of a number of the most well developed cases fails to give us reason to prefer a rich content account.

Recall the issues motivating this metaphysical/empirical investigation which I discussed in the first chapter. For one, if we consider that perceptual content may be limited to colors, shapes, spatial relations, and the like, our attributions of richer properties like pine-tree-ness, beauty, and climbability therefore likely involve a significant conceptual component: “what it’s like” to have a conscious visual experience is exhausted by properties like shape, color, motion, spatial, and illumination properties, and high-level properties enter the picture when conceptual resources are brought to bear, later in processing than conscious visual experience.

These metaphysical results have direct epistemological consequences, as well. Insofar as we lack reason to think there is rich perceptual content, we also lack independent reason for thinking that beliefs about rich content properties like moral and aesthetic properties enjoy immediate perceptual justification. On a thin account of perceptual content the foundationalist’s stock of immediately justified beliefs do include some external world properties – on my view, foundationalists get lots of immediately justified beliefs about surfaces textures, spatial properties, and others. But I’ve argued that, at least presently, we lack reason to think that conscious perceptual experience will directly justify beliefs beyond the thin sort involving surfaces, shapes, colors, and motion: there may be quite a lot of work left for inference to do.

For example, applying the above to the case of beauty, insofar as your visual experience is limited to the low-level properties of the face, you may have a powerful, mandatory mechanism that results in beliefs about beauty. However, I’ve argued that
there’s insufficient ground for thinking this property figures into conscious visual experience. Your experience can immediately justify beliefs about lines and surfaces, but attributions of beauty may require the deployment of conceptual resources after conscious visual experience. Similarly, we lack reason to think perceptual experience will provide an easy route to justified beliefs about other minds through experience of facial expression types (as discussed in connection with Block 2014) or otherwise, leaving theory-theory and simulation theory as the primary competitors. Similarly, I argued that data from Michotte launchings can be well explained without appeal to experience as of mechanical causation, and, accordingly, these data do not give us reason to think that we have immediate justification for our beliefs about mechanical causation. In my view, the Humean challenge to show how we can arrive at justified beliefs about necessary connection still stands.

I argued similarly with respect to affordances, rejecting the Gibsonian claim that perception is primarily of affordances, running counter to the kind of lines found in, among others, Heidegger and Merleau-Ponty according to which our fundamental experiential acquaintance with the world is already imbued with meaning. On the basis of my argument against the Gibsonian claim in Chapter I, we may yet still require, for proper theorizing about our perceptual acquaintance with the world, the kind of “objective/subjective distinction” Gibson was urging us to abandon (depending, of course, on how one interprets Gibson). It may be that our perceptual experience provides epistemic contact with low level, objective, albeit organism-relative, shape and spatial features of the environment, and we, in a certain cognitive sense, post-experientially generate attributions of properties like “too hot for me” or “graspable by me”.

139
Further, on the view I suggested, we need not fear that affordance perception brings the threat of a veil of action-oriented-perception preventing perceptual contact with the “objective” features of the world, thus answering the skeptical worries presented by Kathleen Akins.

In Chapter III, I discuss my second main topic: how visual experience comes to have the content it does, and what effect this might have on its ability to directly justify perceptual beliefs. I argued that the answer to this question is a bit complicated. Cognitive penetration cases bring out two seemingly opposed epistemic intuitions, and I’ve argued that, since neither extinguishes the intuitive force of the other, we should prefer an account that can accommodate two epistemic dimensions of justification, perhaps one deontologically-oriented and one accuracy-oriented, which would explain both intuitions. In order to maintain this position, I defended both reliabilism and phenomenal conservatism from claims that cognitive penetration related intuitions speak decisively against each view. I also provided criticism of mentalist (mostly mental state internalist but access-externalist) competitors to these two accounts, provided in McGrath 2013 and Siegel 2010, and Siegel 2017, finding in each case that the competitor was neither better at accounting for cognitive penetration intuitions nor supported by an argument explaining away the intuitions it could not account for. As a result, cognitive penetration cases push in favor of any account which can accommodate both.

In the process, I engaged at length with the proposal that perceptual experience itself should be understood as having a very belief-like epistemic role, not only subject to potential etiology-based losses in justifying power, but capable of itself being rational and irrational, and redounding on the rational standing of the agent. On this, I sided with
Ernest Sosa (of 2007), arguing that we don’t have sufficient reason to think that there are explanatory gains to be made in extending the “house of reason” to include conscious perceptual experience. I argued that, on the contrary, there are a number of potential disanalogies between belief and experience that may be good reason to think they occupy different epistemic roles.

In summary, the picture that has emerged is the following: first, investigation into high level perceptual content reveals a defensible low-level interpretation as well. We should not, at this point, count on the ability to explain our access to high-level properties in a way that depends upon immediate perceptual justification of our beliefs about those properties. Second, cognitive penetration does show us something about the epistemic role experience plays, but it does not show that perceptual experience shouldn’t have a special epistemic role at all. Rather, it’s still quite plausible that perceptual experience has a special epistemic role, and the intuitions in cognitive penetration cases show two dimensions of epistemic evaluation that a theory, other things being equal, would do well to make room for.
Sources


Carruthers, P. “Who is blind to blindsight?.” Psyche 7 (4) (2001)


Lyons, J. (2005). "Clades, Capgras, and Perceptual Kinds". *Philosophical Topics* 33:185-


Nanay, B. (2011). “Do we see apples as edible?” *Pacific Philosophical Quarterly* 92 (3):305-322


Schwiedrzik, C., Ruff, C. C., et al. (2013). “Untangling Perceptual Memory: Hysteresis and Adaptation Map into Separate Cortical Networks.” *Cerebral Cortex*


