BEHAVIORAL AND BRAIN SCIENCES (2016), Page 1 of 77
doi:10.1017/S0140525X15000965, e229

Cognition does not affect perception: Evaluating the evidence for "top-down" effects

Chaz Firestone
Department of Psychology, Yale University, New Haven, CT 06520-8205
chaz.firestone@yale.edu

Brian J. Scholl
Department of Psychology, Yale University, New Haven, CT 06520-8205
brian.scholl@yale.edu

Abstract: What determines what we see? In contrast to the traditional "modular" understanding of perception, according to which visual processing is encapsulated from higher-level cognition, a tidal wave of recent research alleges that states such as beliefs, desires, emotions, motivations, intentions, and linguistic representations exert direct, top-down influences on what we see. There is a growing consensus that such effects are ubiquitous, and that the distinction between perception and cognition may itself be unsustainable. We argue otherwise: None of these hundreds of studies – either individually or collectively – provides compelling evidence for true top-down effects on perception, or "cognitive penetrability." In particular, and despite their variety, we suggest that these studies all fall prey to only a handful of pitfalls. And whereas abstract theoretical challenges have failed to resolve this debate in the past, our presentation of these pitfalls is empirically anchored: In each case, we show not only how certain studies *could* be susceptible to the pitfall (in principle), but also how several alleged top-down effects *actually are* explained by the pitfall (in practice). Moreover, these pitfalls are perfectly general, with each applying to dozens of other top-down effects. We conclude by extracting the lessons provided by these pitfalls into a checklist that future work could use to convincingly demonstrate top-down effects on visual perception. The discovery of substantive top-down effects of cognition on perception would revolutionize our understanding of how the mind is organized; but without addressing these pitfalls, no such empirical report will license such exciting conclusions.

1. Introduction

How does the mind work? Though this is, of course, the central question posed by cognitive science, one of the deepest insights of the last half-century is that the question does not have a single answer: There is no one way the mind works, because the mind is not one thing. Instead, the mind has parts, and the different parts of the mind operate in different ways: Seeing a color works differently than planning a vacation, which works differently than understanding a sentence, moving a limb, remembering a fact, or feeling an emotion.

The challenge of understanding the natural world is to capture generalizations – to "carve nature at its joints." Where are the joints of the mind? Easily, the most natural and robust distinction between types of mental processes is that between perception and cognition. This distinction is woven so deeply into cognitive science as to structure introductory courses and textbooks, differentiate scholarly journals, and organize academic departments. It is also a distinction respected by common sense: Anyone can appreciate the difference between, on the one hand, *seeing* a red apple and, on the other hand, *thinking about, remembering,* or *desiring* a red apple. This difference is especially clear when perception and cognition deliver conflicting evidence about the world – as in most visual illusions. Indeed, there may be no better way to truly *feel* the distinction between perception and cognition for yourself than to visually experience the world in a way you know it not to be.

There is a deep sense in which we all know what perception is because of our direct phenomenological acquaintance with *percepts* – the colors, shapes, and sizes (etc.) of the objects and surfaces that populate our visual experiences. Just imagine looking at an apple in a supermarket and appreciating its redness (as opposed, say, to its price). *That* is perception. Or look at Figure 1A and notice the difference in lightness between the two gray rectangles. *That* is perception. Throughout this paper, we refer to *visual processing* simply as the mental activity that creates such sensations; we refer to *percepts* as the experiences themselves, and we use *perception* (and, less formally, *seeing*) to encompass both (typically unconscious) visual processing and the (conscious) percepts that result.

1.1. The new top-down challenge

Despite the explanatorily powerful and deeply intuitive nature of the distinction between seeing and thinking, a

0140-525X/16

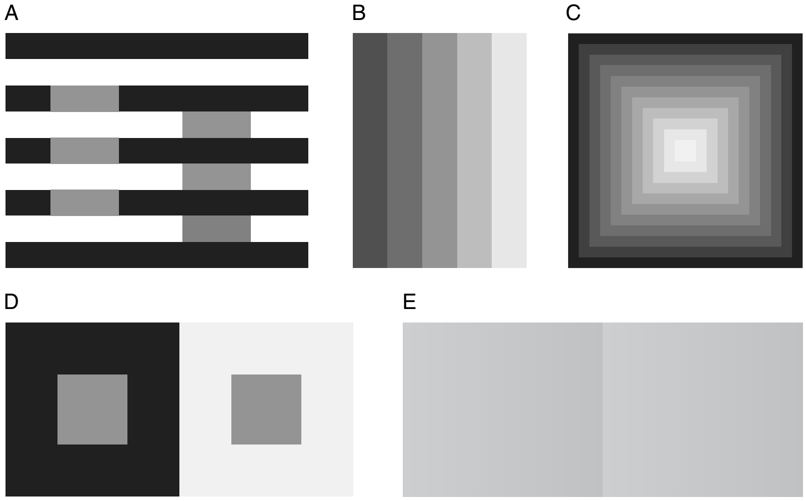


Figure 1. Examples of lightness illusions can be subjectively appreciated as "demonstrations" (for references and explanations, see Adelson 2000). (**A**) The two columns of gray rectangles have the same luminance, but the left one looks lighter. (**B**) The rectangles are uniformly gray, but they appear to lighten and darken along their edges. (**C**) Uniformly colored squares of increasing luminance produce an illusory light "X" shape at their corners. (**D**) The two central squares have the same objective luminance, but the left one looks lighter. (**E**) The two rectangles are identical segments of the same gradient, but the right one looks lighter. Similar demonstrations abound, for nearly every visual feature.

vocal chorus has recently and vigorously challenged the extent of this division, calling for a generous blurring of the lines between visual perception and cognition (for recent reviews, see Balcetis 2016; Collins & Olson 2014; Dunning & Balcetis 2013; Goldstone et al. 2015; Lupyan 2012; Proffitt & Linkenauger 2013; Riccio et al. 2013; Stefanucci et al. 2011; Vetter & Newen 2014; Zadra & Clore 2011). On this increasingly popular view, higher-level cognitive states routinely "penetrate" perception, such that what we see is an alloy both of bottom-up factors and of beliefs, desires, motivations, linguistic representations, and other such states. In other words, these views hold that the mental processes responsible for building percepts can and do access radically more information elsewhere in the mind than has traditionally been imagined.

Chaz Firestone is a graduate student in the Department of Psychology at Yale University. As of July 2017, he will be an Assistant Professor in the Department of Psychological and Brain Sciences at Johns Hopkins University. He holds an Sc.B. in cognitive neuroscience and an A.M. in philosophy, both from Brown University. His research explores the border between perception and cognition, and he was recognized for this work with the 2013 William James Prize from the Society for Philosophy and Psychology. He hasn't published very many papers, but this one is his favorite.

Brian Scholl is Professor of Psychology and Chair of the Cognitive Science program at Yale University, where he also directs the Perception & Cognition Laboratory. He and his research group have published more than 100 papers on various topics in cognitive science, with a special focus on how visual perception interacts with the rest of the mind. He is a recipient of the Distinguished Scientific Award for Early Career Contribution to Psychology and the Robert L. Fantz Memorial Award, both from the American Psychological Association, and is a past President of the Society for Philosophy and Psychology. At Yale he is a recipient of both the Graduate Mentor Award and the Lex Hixon Prize for Teaching Excellence in the Social Sciences, and he has great fun teaching the Introduction to Cognitive Science course.

At the center of this dispute over the nature of visual perception and its relation to other processes in the mind has been the recent and vigorous proliferation of so-called top-down effects on perception. In such cases, some extraperceptual state is said to literally and directly alter what we see. (As of this writing, we count more than 175 papers published since 1995 reporting such effects; for a list, see http://perception.yale.edu/TopDownPapers.) For example, it has been reported that desiring an object makes it look closer (Balcetis & Dunning 2010), that reflecting on unethical actions makes the world look darker (Banerjee et al. 2012), that wearing a heavy backpack makes hills look steeper (Bhalla & Proffitt 1999), that words having to do with morality are easier to see (Gantman & Van Bavel 2014), and that racial categorization alters the perceived lightness of faces (Levin & Banaji 2006).

If what we think, desire, or intend (etc.) can affect what we see in these ways, then a genuine revolution in our understanding of perception is in order. Notice, for example, that the vast majority of models in vision science do not consider such factors; yet, apparently, such models have been successful! For example, today's vision science has essentially worked out how low-level complex motion is perceived and processed by the brain, with elegant models of such processes accounting for extraordinary proportions of variance in motion processing (e.g., Rust et al. 2006) – and this success has come without factoring in morality, hunger, or language (etc.). Similarly, such factors are entirely missing from contemporary vision science textbooks (e.g., Blake & Sekuler 2005; Howard & Rogers 2002; Yantis 2013). If

http://dx.doi.org/10.1017/S0140525X15000965

such factors do influence how we see, then such models and textbooks are scandalously incomplete.

Although such factors as morality, hunger, and language are largely absent from contemporary vision science in practice, the emergence of so many empirical papers reporting top-down effects of cognition on perception has shifted the broader consensus in cognitive science. Indeed, such alleged top-down effects have led several authors to declare that the revolution in our understanding of perception has already occurred, proclaiming as dead not only a "modular" perspective on vision, but often also the very distinction between perception and cognition itself. For example, it has been asserted that it is a "generally accepted concept that people tend to see what they want to see" (Radel & Clément-Guillotin 2012, p. 233), and that "the postulation of the existence of visual processes being functionally encapsulated…cannot be justified anymore" (Vetter & Newen 2014, p. 73). This sort of evidence led one philosopher to declare, in an especially sweeping claim, that "[a]ll this makes the lines between perception and cognition fuzzy, perhaps even vanishing" and to deny that there is "any real distinction between perception and belief" (Clark 2013, p. 190).

1.2. Our thesis and approach

Against this wealth of evidence and its associated consensus, the thesis of this paper is that there is in fact *no* evidence for such top-down effects of cognition on visual perception, in every sense these claims intend. With hundreds of reported top-down effects, this is, admittedly, an ambitious claim. Our aim in this discussion is thus to explicitly identify the (surprisingly few, and theoretically interesting) "pitfalls" that account for reports of top-down penetration of visual perception without licensing such conclusions.

Our project differs from previous theoretical challenges (e.g., Fodor 1984; Pylyshyn 1999; Raftopoulos 2001a) in several ways. First, whereas many previous discussions defended the modular nature of only a circumscribed (and possibly unconscious) portion of visual processing (e.g., "early vision"; Pylyshyn 1999), we have the broader aim of evaluating the evidence for top-down effects on *what we see* as a whole – including visual processing and the conscious percepts it produces. Second, several pitfalls we present are novel contributions to this debate. Third, and most important, whereas past abstract discussions have failed to resolve this debate, our presentation of these pitfalls is empirically anchored: In each case, we show not only how certain studies *could* be susceptible to the pitfall (in principle), but also how several alleged top-down effects *actually are* explained by the pitfall (in practice, drawing on recent and decisive empirical studies). Moreover, each pitfall we present is perfectly general, applying to dozens more reported top-down effects. Research on top-down effects on visual perception must therefore take the pitfalls seriously before claims of such phenomena can be compelling.

The question of whether there are top-down effects of cognition on visual perception is one of the most foundational questions that can be asked about what perception is and how it works, and it is therefore no surprise that the issue has been of tremendous interest (especially recently) – not only in all corners of psychology, but also in neighboring disciplines such as philosophy of mind (e.g., Macpherson 2012; Siegel 2012), neuroscience (e.g., Bannert & Bartels 2013; Landau et al. 2010), psychiatry (e.g., Bubl et al. 2010), and even aesthetics (e.g., Nanay 2014; Stokes 2014). It would be enormously exciting to discover that perception changes the way it operates in direct response to goings-on elsewhere in the mind. Our hope is thus to help advance future work on this foundational question, by identifying and highlighting the key empirical challenges.

2. A recipe for revolution

The term *top-down* is used in a spectacular variety of ways across many literatures. What do we mean when we say that cognition does not affect perception, such that there are no top-down effects on what we see? The primary reason these issues have received so much historical and contemporary attention is that a proper understanding of mental organization depends on whether there is a salient "joint" between perception and cognition. Accordingly, we focus on the sense of *top-down* that directly addresses this aspect of how the mind is organized. This sense of the term is, for us, related to traditional questions of whether visual perception is modular, encapsulated from the rest of cognition, and "cognitively (im)penetrable."[1] At issue is the extent to which what and how we *see* is functionally independent from what and how we think, know, desire, act, and so forth. We single out this meaning of *top-down* not only because it may be the most prominent usage of the term, but also because the questions it raises are especially foundational for our understanding of the organization of the mind.

Nevertheless, there are several independent uses of *top-down* that are less revolutionary and do not directly interact with these questions.

2.1. Changing the processing versus (merely) changing the input

On an especially permissive reading of "top-down," top-down effects are all around us, and it would be absurd to deny cognitive effects on what we see. For example, there is a trivial sense in which we all can willfully control what we visually experience, by (say) choosing to close our eyes (or turn off the lights) if we wish to experience darkness. Though this is certainly a case of cognition (specifically, of desire and intention) changing perception, this familiar "top-down" effect clearly isn't revolutionary, insofar as it has no implications for how the mind is organized – and for an obvious reason: Closing your eyes (or turning off the lights) changes only the *input* to perception; it does not change perceptual processing itself.

Despite the triviality of this example, the distinction is worth keeping in mind, because it is not always obvious when an effect operates by changing the input. To take one fascinating example, facial expressions associated with fear (e.g., widened eyes) and disgust (e.g., narrowed eyes) have recently been shown to reliably vary the eye-aperture diameter, directly influencing acuity and sensitivity by altering the actual optical information reaching perceptual processing (Lee et al. 2014). (As we will see later, the distinction between input and processing also arises with regard to perceptual vs. attentional effects.)

2.2. Descending neural pathways

In systems neuroscience, some early models of brain function were largely feedforward, with various brain regions feeding information to each other in a unidirectional sequence. In contrast, there is now considerable evidence that brain regions that were initially considered "higher up" in a processing hierarchy can modulate "lower" regions, through so-called re-entrant processing from descending neural pathways – and these sorts of modulation are often also commonly called top-down effects (e.g., Gilbert & Li 2013; Rolls 2008; Zhang et al. 2014). Though extremely interesting for certain questions about functional neuroanatomy, this type of "top-down" influence has no necessary implications for cognitive penetrability. One reason is that nearly all brain regions subserve multiple functions. Even parts of visual cortex, for example, are involved in imagery (e.g., Kosslyn 2005), recall (e.g., Le Bihan et al. 1993), and reward processing (Vickery et al. 2011) – so that it is almost never clear which mental process a descending pathway is descending *to* (or if that descending pathway is influencing the input or the processing of whatever it descends to, per sect. 2.1).

At any rate, we do not discuss descending pathways in the brain in this target article, for two reasons. First, the implications of this body of work for issues of modularity and cognitive penetrability have been addressed and critiqued extensively elsewhere (e.g., Raftopoulos 2001b). Second, our aim here is to focus on that recent wave of work that promises a revolution in how we think about the organization of the mind. And whatever one thinks of the relevance of descending neural pathways to issues of whether cognition affects perception, they certainly cannot be revolutionary today: The existence of descending neural pathways has been conclusively established many times over, and they are now firmly part of the orthodoxy in our understanding of neural systems.

2.3. Top-down effects versus context effects and "unconscious inferences" in vision

Visual processing is often said to involve "problem solving" (Rock 1983) or "unconscious inference" (Gregory 1980; Helmholtz 1866/1925). Sometimes these labels are applied to seemingly sophisticated processing, as in research on the perception of causality (e.g., Rolfs et al. 2013; Scholl & Tremoulet 2000) or animacy (e.g., Gao et al. 2010; Scholl & Gao 2013). But more often, the labels are applied to relatively early and low-level visual processing, as in the perception of lightness (e.g., Adelson 2000) or depth (e.g., Ramachandran 1988). In those cases, such terminology (which may otherwise evoke notions of cognitive penetrability) refers to aspects of processing that are wired into the visual module itself (so-called "natural constraints") – and so do not at all imply effects of cognition on perception, or "top-down" effects. This is true even when such processing involves *context effects*, wherein perception of an object may be influenced by properties of other objects nearby (e.g., as in several of the lightness illusions in Fig. 1). In such cases, the underlying processes continue to operate reflexively (based solely on their visual input) regardless of *your* cognitive inferences or problem-solving strategies (for discussion, see Scholl & Gao 2013) – as when lightness illusions based on "unconscious inferences" persist in the face of countervailing knowledge (Fig. 1). (For further discussion of why vision being "smart" in such ways does not imply cognitive penetrability, see Kanizsa 1985; Pylyshyn 1999.)

2.4. Cross-modal effects

What we see is sometimes affected by other sense modalities. For example, a single flash of light can appear to flicker when accompanied by multiple auditory beeps (Shams et al. 2000), and two moving discs that momentarily overlap are seen to bounce off each other (rather than stream past each other) if a beep is heard at the moment of overlap (Sekuler et al. 1997). However, these cases – though interesting for many other reasons – do not demonstrate cognitive penetrability, for much the same reason that unconscious inferences in vision fail to do so. For example, such crossmodal integration is itself a reflexive, apparently impenetrable process: The sounds' effects occur "whether you like it or not," and they occur extremely quickly (e.g., in less than 100 ms; Shams et al. 2002). Collectively, such results are consistent with the entire process being contained within perception itself, rather than being an effect of more central cognitive processes *on* perception.

At any rate, we do not discuss crossmodal effects here. As with descending neural pathways, whatever one thinks of the relevance of this work to the issues we discuss, they certainly cannot be revolutionary today in the way promised by the work we review in section 3 – if only because the existence of crossmodal effects has been conclusively established and is common ground for all parties in this discussion.

2.5. Input-driven changes in sensitivity over time

Despite encapsulation, input may sometimes change visual processing by increasing sensitivity over time to certain visual features. For example, figure–ground assignment for ambiguous stimuli is sometimes biased by experience: The visual system will more likely assign *figure* to familiar shapes, such as the profile of a woman with a skirt (Peterson & Gibson 1993; Fig. 2A). However, such changes don't involve any penetration because they don't involve effects of knowledge per se. For example, inversion eliminates this effect even when subjects know the inverted shape's identity (Peterson & Gibson 1994). Therefore, what may superficially appear to be an influence of knowledge on perception is simply increased sensitivity to certain contours. Indeed, Peterson and Gibson (1994) volunteer that their phenomena don't reflect top-down effects, and in particular that "the orientation dependence of our results demonstrates that our phenomena are not dependent on semantic knowledge" (p. 561). Thus, such effects aren't "top-down" in any sense that implies cognitive penetrability, because the would-be penetrator is just the low-level visual input itself. (Put more generally, the thesis of cognitive impenetrability constrains the information modules can access, but it does not constrain what modules can do with the input they *do* receive; e.g., Scholl & Leslie 1999.)

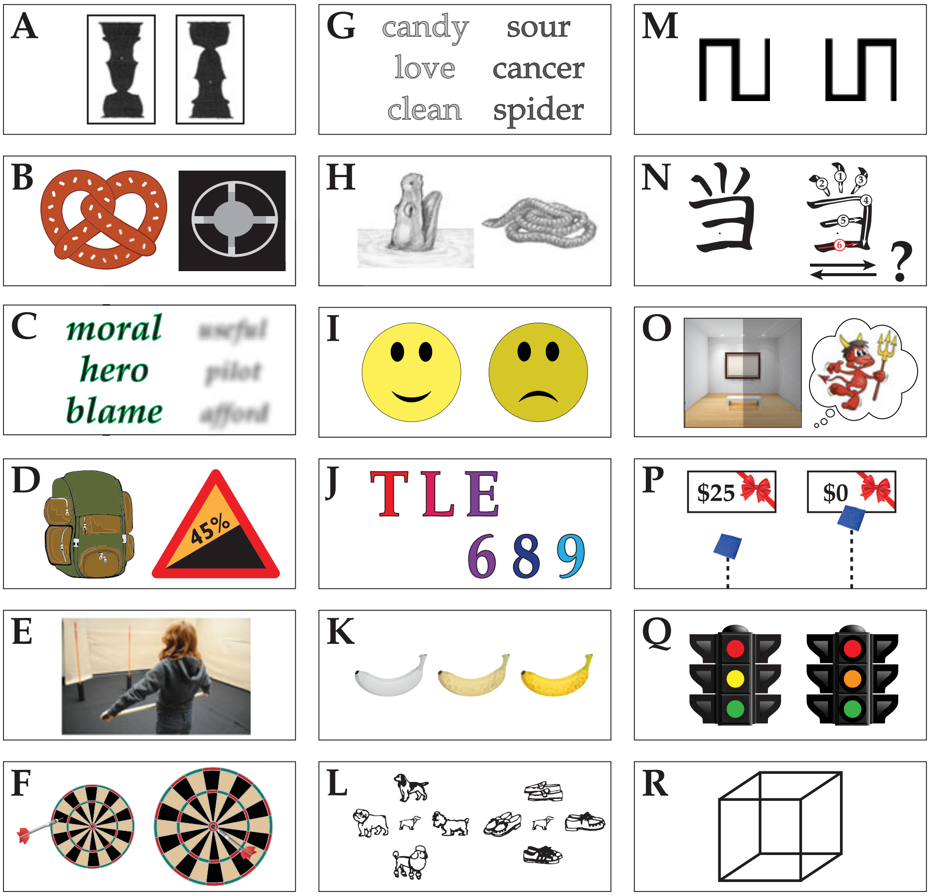


Figure 2. Diagrams or depictions of various possible top-down effects on perception: (**A**) Figure–ground assignment is biased toward familiar shapes, such as the profile of a woman (Peterson & Gibson 1993). (**B**) Being thirsty (as a result of eating salty pretzels) makes ambiguous surfaces look more transparent (Changizi & Hall 2001). (**C**) Morally relevant words are easier to see than morally irrelevant words (Gantman & Van Bavel 2014). (**D**) Wearing a heavy backpack makes hills look steeper (Bhalla & Proffitt 1999). (**E**) Holding a wide pole makes apertures look narrower (Stefanucci & Geuss 2009). (**F**) Accuracy in dart throwing biases subsequent estimates of the target's size (Cañal-Bruland et al. 2010; Wesp & Gasper 2004). (**G**) Positive words are seen as lighter than negative words (Meier et al. 2007). (**H**) Scary music makes ambiguous images take on their scarier interpretation (Prinz & Seidel 2012). (**I**) Smiling faces appear brighter (Song et al. 2012). (**J**) Learning color–letter associations makes identically hued numbers and letters appear to have their categories' hues (e.g., the E will look red and the 6 will look blue, even though they are equally violet; Goldstone 1995). (**K**) A grayscale banana appears yellow (Hansen et al. 2006). (**L**) Conceptual similarity enhances size-contrast illusions (Coren & Enns 1993). (**M**) Labeling certain blocky figures as "2" and "5" makes them easier to find in a visual search array (Lupyan & Spivey 2008). (**N**) Calligraphic knowledge (e.g., of the direction of the sixth stroke of a Chinese character) affects the direction of apparent motion when that stroke is flashed (Tse & Cavanagh 2000). (**O**) Reflecting on unethical actions makes the world look darker (Banerjee et al. 2012). (**P**) Desired objects are seen as closer, as measured by beanbag throws (Balcetis & Dunning 2010). (**Q**) The middle traffic light is called *gelb* (yellow) in German and *oranje* (orange) in Dutch, which influences its perceived color (Mitterer et al. 2009). (**R**) You may be able to intentionally decide which interpretation of a Necker cube to see (cf. Long & Toppino 2004).

3. Contemporary top-down effects

What remains after setting aside alternative meanings of "top-down effects" is the provocative claim that our beliefs, desires, emotions, actions, and even the languages we speak can directly influence what we see. Much ink has been spilled arguing whether this should or shouldn't be true, based primarily on various theoretical considerations (e.g., Churchland 1988; Churchland et al. 1994; Firestone 2013a; Fodor 1983; 1984; 1988; Goldstone & Barsalou 1998; Lupyan 2012; Machery 2015; Proffitt & Linkenauger 2013; Pylyshyn 1999; Raftopoulos 2001b; Vetter & Newen 2014; Zeimbekis & Raftopoulos 2015). We will not engage those arguments directly – largely, we admit, out of pessimism that such arguments can be (or have been) decisive. Instead, our focus will be on the nature and strength of the empirical evidence for cognitive penetrability in practice.

Though recent years have seen an unparalleled proliferation of alleged top-down effects, such demonstrations have a long and storied history. One especially visible landmark in this respect was the publication in 1947 of Bruner and Goodman's "Value and need as organizing factors in perception." Bruner and Goodman's pioneering study reported that children perceived coins as larger than they perceived worthless cardboard discs of the same physical size, and also that children from poor families perceived the coins as larger than did wealthy children. These early results ignited the New Look movement in perceptual psychology, triggering countless studies purporting to show all manner of top-down influences on perception (for a review, see Bruner 1957). It was claimed, for example, that hunger biased the visual interpretation of ambiguous images (Lazarus et al. 1953), that knowledge of objects' typical colors influenced online color perception (Bruner et al. 1951), and that meaningful religious iconography dominated other symbols in binocular rivalry (Lo Sciuto & Hartley 1963).

However, the New Look movement's momentum eventually stalled as its findings buckled under methodological and theoretical scrutiny. For example, follow-up studies on the value-based size-distortion effects could replicate them only when subjects made judgments from memory rather than during online viewing (Carter & Schooler 1949; see also Landis et al. 1966), and other critiques identified theoretically puzzling moderating variables or reported that many other valuable objects and symbols failed to produce similar results (e.g., Klein et al. 1951; McCurdy 1956). Other confounding variables were eventually implicated in the original effects, leading several researchers to conclude that "[o]nly when better experiments have been carried out will we be able to determine what portion of the effect is due to nonperceptual factors" (Landis et al. 1966, p. 729). By the next decade, the excitement surrounding such ideas had fizzled, and "the word 'artifact' became the descriptive term par excellence associated with the New Look" (Erdelyi 1974, p. 2).

The last two decades have seen the pendulum swing again, away from a robust division between perceptual and cognitive processing and back toward the previously fashionable New Look understanding of perception. The driving force in recent years has been a tidal wave of studies seeming to show influences on perception from all corners of the mind. However, the particular theoretical motivations behind these various results are nonuniform, so it will be useful to understand these studies in groups. Roughly, today's alleged top-down effects on perception are effects of motivation, action, emotion, categorization, and language.

3.1. Motivation

Those recent results with the greatest overlap with the New Look movement concern influences of motivation (desires, needs, values, etc.) on perception. For example, it has recently been reported that desirable objects such as chocolate look closer than undesirable objects such as feces (Balcetis & Dunning 2010; see also Krpan & Schnall 2014); that rewarding subjects for seeing certain interpretations of ambiguous visual stimuli actually makes the stimuli look that way (Balcetis & Dunning 2006; see also Pascucci & Turatto 2013); that desirable destinations seem closer than undesirable ones (Alter & Balcetis 2011; see also Balcetis et al. 2012); and even that women's breasts appear larger to sex-primed men (den Daas et al. 2013). Other studies have focused on physiological needs. For example, muffins are judged as larger by dieting subjects (van Koningsbruggen et al. 2011), food-related words are easier to identify when observers are fasting (Radel & Clément-Guillotin 2012), and ambiguous surfaces are judged as more transparent (or "water-like") by subjects who eat salty pretzels and become thirsty (Changizi & Hall 2001; Fig. 2B). Morally relevant words reportedly "pop out" in visual awareness when briefly presented (Gantman & Van Bavel 2014; Fig. 2C), and follow-up studies suggest that the effect may arise from a desire for justice. Many of these contemporary studies explicitly take inspiration from the New Look, claiming to study the same phenomena but "armed with improved methodological tools and theories" (Dunning & Balcetis 2013, p. 33).

3.2. Action

Another class of recent top-down effects concerns action-based influences on perception. Physical burdens that make actions more difficult reportedly make the environment look more imposing: wearing a heavy backpack inflates estimates of distance (Proffitt et al. 2003), as does throwing a heavy ball (Witt et al. 2004); fatigued or unfit individuals overestimate slant and distance relative to rested or fit individuals (Bhalla & Proffitt 1999; Cole et al. 2013; Sugovic & Witt 2013; Fig. 2D); fixing weights to subjects' ankles increases size estimates of jumpable gaps (Lessard et al. 2009); holding one's arms out decreases width estimates of doorway-like apertures (Stefanucci & Geuss 2009; Fig. 2E); and standing on a wobbly balancing board reduces width estimates of a walkable beam (Geuss et al. 2010). Conversely, improvements in ability are reported to shrink the perceived environment to make actions look easier: Subjects who hold reach-extending batons judge targets to be closer (Witt et al. 2005; see also Abrams & Weidler 2015); subjects who drink a sugary beverage (rather than a low-calorie alternative) estimate hills as shallower (Schnall et al. 2010); and swimmers who wear flippers judge underwater targets as closer (Witt et al. 2011). Similarly, exceptional athletic performance is reported to alter the perceived size of various types of sporting equipment, yielding perceptual reports of larger softballs (Gray 2013; Witt & Proffitt 2005), wider football goal posts (Witt & Dorsch 2009), lower tennis nets (Witt & Sugovic 2010), larger dartboards (Cañal-Bruland et al. 2010; Wesp et al. 2004; Fig. 2F), larger golf holes (Witt et al. 2008), and (for parkour experts) shorter walls (Taylor et al. 2011). This approach emphasizes the primacy of action in perception (inspired in many ways by Gibson 1979), holding that action capabilities directly alter the perceived environment (for reviews, see Proffitt 2006; Proffitt & Linkenauger 2013; Witt 2011a). (Though it is not entirely clear whether action per se is a truly *cognitive* process, we mean to defend an extremely broad thesis regarding the sorts of states that cannot affect perception, and this most definitely includes action. Moreover, in many of these cases, it has been proposed that it is not the action that penetrates perception but rather the *intention* to act – e.g., Witt et al. 2005 – in which case such effects would count as alleged examples of cognition affecting perception after all.)

3.3. Affect and emotion

A third broad category of recently reported top-down effects involves affective and emotional states. In such cases, the perceived environment is purportedly altered to match the perceiver's mood or feelings. For example, recent studies report that thinking negative thoughts makes the world look darker (Banerjee et al. 2012; Meier et al. 2007; Fig. 2G); fear and negative arousal make hills look steeper, heights look higher, and objects look closer (Cole et al. 2012; Harber et al. 2011; Riener et al. 2011; Stefanucci & Proffitt 2009; Stefanucci & Storbeck 2009;

Stefanucci et al. 2012; Storbeck & Stefanucci 2014; Teachman et al. 2008); scary music makes ambiguous images (e.g., an ambiguous figure that might be an alligator or a squirrel) take on their scarier interpretations (Prinz & Seidel 2012; Fig. 2H); social exclusion makes other people look closer (Pitts et al. 2014); and smiling faces appear brighter (Song et al. 2012; Fig. 2I). Here, the effects are either thought to accentuate one's emotional state – perhaps because affect is informative about the organism's needs (e.g., Storbeck & Clore 2008) – or to energize the perceiver to counteract such negative feelings.

3.4. Categorization and language

A final class of contemporary top-down effects concerns categories and linguistic labels. A popular testing ground for such effects has involved the perception of color and lightness. For example, it has been reported that learning color–letter associations biases perceptual judgments toward the learned hues (Goldstone 1995; Fig. 2J); categorizing faces as Black or White alters the faces' perceived skin tones, even when the faces are in fact equally luminant (Levin & Banaji 2006); and knowledge of an object's typical color (e.g., that bananas are yellow) makes grayscale images of those objects appear tinged with their typical colors (Hansen et al. 2006; Witzel et al. 2011; e.g., Fig. 2K). Conceptual categorization is also reported to modulate various visual phenomena. For example, the Ebbinghaus illusion, in which a central image appears smaller when surrounded by large images (or larger when surrounded by small images), is reportedly stronger when the surrounding images belong to the same conceptual category as the central image (Fig. 2L; Coren & Enns 1993; see also van Ulzen et al. 2008).

Similar effects may arise from linguistic categories and labels. For example, the use of particular color terms is reported to affect how colors actually appear (e.g., Webster & Kay 2012), and labeling visual stimuli reportedly enhances processing of such stimuli and may even alter their appearance (Lupyan & Spivey 2008; Lupyan et al. 2010; Lupyan & Ward 2013; Fig. 2M). Other alleged linguistic effects include reports of visual motion aftereffects after reading motion-related language (e.g., "Google's stock sinks lower than ever"; Dils & Boroditsky 2010a; 2010b; see also Meteyard et al. 2007), and differences in the apparent motion of a Chinese character's stroke depending on knowledge of how such characters are written (Tse & Cavanagh 2000; though see Li & Yeh 2003; Fig. 2N).

Note that the effects cited in this section are not only numerous and varied, but also they are exceptionally *recent*: Indeed, the median publication year for the empirical papers cited in section 3 is 2010.

4. The six "pitfalls" of top-down effects on perception

If there are no top-down effects of cognition on perception, then how have so many studies seemed to find such rich and varied evidence for them? A primary purpose of this paper is to account for the wealth of research reporting such top-down effects. We suggest that this research falls prey to a set of "pitfalls" that undermine their claims. These pitfalls have four primary features:

1. *They are few in number.* We suggest that nearly all of the recent literature on top-down effects is susceptible to a surprisingly small group of such pitfalls.
2. *They are empirically anchored.* These are not idle suspicions about potential causes of such effects, but rather they are empirically grounded – not just in the weak sense that they discuss relevant empirical evidence, but in the stronger sense that they have demonstrably explained several of the most prominent apparent top-down effects on perception, in practice.
3. *They are general in scope.* Beyond our concrete case studies, we also aim to show that the pitfalls are broadly applicable, with each covering dozens more top-down effects.
4. *They are theoretically rich.* Exploring these pitfalls raises several foundational questions not just about perception and cognition, but also about their relationships with other mental processes, including memory, attention, and judgment.

We contend that any apparent top-down effect that falls prey to one or more of these pitfalls would be compromised, in the sense that it could be explained by deflationary, routine, and certainly nonrevolutionary factors. It is thus our goal to establish the empirical concreteness and general applicability of these pitfalls, so that it is clear where the burden of proof lies: No claim of a top-down effect on perception can be accepted until these pitfalls have been addressed.

We first discuss each pitfall in general terms and then provide empirical case studies of how it can be explored in practice, along with suggestions of other top-down effects to which it may apply. In each case, we conclude with concrete lessons for future research.

4.1. Pitfall 1: An overly confirmatory research strategy

In general, experimental hypotheses can be tested in two sorts of ways: Not only should you observe an effect when your theory calls for it, but also you should *not* observe an effect when your theory demands its absence. Although both kinds of evidence can be decisive, the vast majority of reported top-down effects on perception involve only the first sort of test: a hypothesis is proffered that some higher-level state affects what we see, and then such an effect is observed. Though it is perhaps unsurprising that these studies only test such "confirmatory predictions," in our view this strategy essentially misses out on half of the possible decisive evidence. Recently, this theoretical perspective has been made empirically concrete by studies testing certain kinds of *uniquely disconfirmatory predictions* of various top-down phenomena.

4.1.1. Case studies. To make the contrast between confirmatory and disconfirmatory predictions concrete, we conducted a series of studies (Firestone & Scholl 2014b) inspired by an infamous art-historical reasoning error known as the "El Greco fallacy." Beyond appreciating the virtuosity of his work, the art-history community has long puzzled over the oddly elongated human figures in El Greco's paintings. To explain these distortions, it was once supposed that El Greco suffered from an uncommonly severe astigmatism that effectively "stretched" his perceived environment, such that El Greco had simply been painting what he saw. This perspective was once taken seriously, but upon reflection it involves a conceptual confusion: If El Greco had truly experienced a stretched-out world, then he would also have experienced an equally stretched-out *canvas,* canceling out the supposed real-world distortions and thus leaving no trace of them in his reproductions. The distortions in El Greco's paintings, then, could not reflect literal perceptual distortions (Anstis 2002; Firestone 2013b).

We exploited the El Greco fallacy to show that multiple alleged top-down effects cannot genuinely be effects on perception. Consider, for example, the report that reflecting on unethical actions makes the world look darker (Banerjee et al. 2012; Fig. 2O). The original effect was obtained using a numerical scale: After reflecting on ethical or unethical actions, subjects picked a number on the scale to rate the brightness of the room they were in. We replicated this effect with one small change: Instead of a numerical scale, subjects used a scale of actual grayscale patches to rate the room's brightness. According to the view that reflecting on negative actions makes the world look darker, this small change drastically alters the study's prediction: If the world really looks darker, then the patches making up the scale should look darker too, and the effects should thus cancel each other out (just as the alleged distortions in El Greco's experience of the world would be canceled out by his equally distorted experience of his canvas). However, the follow-up study *succeeded*: subjects still picked a darker patch to match the room after reflecting on an unethical action (Firestone & Scholl 2014b, Experiment 5). This effect, then – like the distortions in El Greco's work – must not reflect the way the world actually looked to subjects.

This approach is in no way limited to the particulars of the morality/brightness study. Indeed, to apply the same logic more broadly, we also explored a report of a very different higher-level state (a subject's ability to act in a certain way) on a very different visual property (perceived distance). In particular, holding a wide rod across one's body (Fig. 2E) reportedly makes the distance between two poles (which form a doorway-like aperture) look narrower, as measured by having subjects instruct the experimenter to adjust a measuring tape to visually match the aperture's width. The effect supposedly arises because holding the rod makes apertures less passable (Stefanucci & Geuss 2009). We successfully replicated this result, but we also tested it with one critical difference: Instead of adjusting a measuring tape to record subjects' width estimates, the experimenter used two poles that themselves formed an independent and potentially passable aperture. Again, the El Greco logic applies: If holding a rod really does perceptually compress apertures, then this variant should "fail," because subjects should see *both* apertures as narrower. But the experiment did not "fail": Subjects again reported narrower apertures even when responding with an aperture (Firestone & Scholl 2014b, Experiment 2). Therefore, this effect cannot reflect a true perceptual distortion – not because the effect fails to occur, but rather because it occurs even when it shouldn't. (In later experiments, we determined the true, nonperceptual, explanation for this effect, involving task demands; see Pitfall 3.)

4.1.2. Other susceptible studies. As an example of testing disconfirmatory predictions, the El Greco fallacy applies to any constant-error distortion that should affect equally the means of reproduction (e.g., canvases, grayscale patches) and the item reproduced (e.g., visual scenes to be painted, bright rooms). The studies that fail to test such predictions are too numerous to count; essentially, nearly every study falls into this category. However, some studies of top-down effects on perception may have tested such predictions inadvertently – and, given their results, perhaps committed the El Greco fallacy.

Consider, for example, the report that after repeatedly viewing one set of red and violet letters and a second set of blue and violet numbers, subjects judged token violet letters to look redder than they truly were and token violet numbers to look bluer than they truly were (Goldstone 1995; Fig. 2J). This effect was measured by having subjects adjust the hue of a stimulus to perceptually match the symbol being tested. However, the adjusted stimulus was a copy of that symbol! For example, after viewing a red "T," a reddish-violet "L," and a violet "E," subjects judged the E to be redder – as measured by adjusting the hue of a second E. This commits the El Greco fallacy: if Es really look redder after one sees other red letters, then both the to-be-matched E and the adjustable E should have looked redder, and the effects should have canceled one another out. That such an effect was nevertheless obtained suggests it cannot be perceptual.

Similarly, consider the following pair of results, reported together: Subjects judged gray patches to be darker after reading negative (vs. positive) words, and subjects judged words printed in gray ink to be darker if the words were negative (vs. positive), as measured by selecting a darker grayscale patch to match the word's lightness (Meier et al. 2007; Fig. 2G). Here too is an El Greco fallacy: if, per the first result, reading negative words makes gray patches look darker, then the gray patches from the second result should also have looked darker, and the effects of one should have canceled out the other.

The El Greco fallacy may also afflict reports that linguistic color categories alter color appearance (Webster & Kay 2012). For example, a color that is objectively between blue and green may appear *either* blue *or* green because our color terms single out those colors when they discretize color space, creating clusters of perceptual similarity. However, such studies use color spaces specifically constructed for perceptual uniformity, such that each step through the space's parameters is perceived as equal in magnitude. This raises a puzzle: If color terms affect perceived color, then such effects should already have been assimilated into the color space, leaving no room for color terms to exert their influence in studies using such color spaces. That these studies still show labeling effects suggests an alternative explanation.[2]

4.1.3. A lesson for future research. To best determine the extent to which cognition influences perception, future studies should proactively employ both confirmatory and disconfirmatory research strategies; to do otherwise is to ignore half of the predictions the relevant theories generate. In pursuing disconfirmatory evidence, El Greco–

inspired research strategies in particular have three distinct advantages. First, they can rule out perceptual explanations without relying on null effects and their attendant interpretive problems; instead, this strategy can disconfirm top-down interpretations through positive replications. Second, the El Greco strategy can fuel such implications even before researchers determine the actual (nonperceptual) culprit (just as we know that astigmatism does not explain El Greco's distortions, even if we remain uncertain what does explain them). Finally, this strategy is broadly relevant – being applicable any time a scale can be influenced just as the critical stimuli are supposedly influenced (e.g., in nearly all perceptual matching tasks).

4.2. *Pitfall 2: Perception versus judgment*

Many alleged top-down effects on perception live near the border of perception and cognition, where it is not always obvious whether a given cognitive state affects what we *see* or instead only our inferences or judgments made *on the basis of* what we see. This distinction is intuitive elsewhere. For example, whereas we can perceive the color or size of some object – say, a shoe – we can only *infer* or *judge* that the object is expensive, comfortable, or fashionable (even if we do so based on how it looks). Top-down effects on perception pose a special interpretive challenge along these lines, especially when they rely on subjects' verbal reports. Whereas expensiveness can *only* be judged (not perceived), other properties such as color and size can be both perceived and judged: We can directly see that an object is red, and we can also conclude or infer that an object is red. For this reason, any time an experiment shifts perceptual reports, it is possible that the shift reflects changes in judgment rather than perception. And whereas top-down effects on perception would be revolutionary and consequential, many top-down effects on judgments are routine and unsurprising, carrying few implications for the organization of the mind. (Of course, that is not to say that research on judgment in general is not often of great interest and import – just that some effects on judgment truly are routine and universally accepted, and those are the ones that may explain away certain purported top-down effects of cognition on perception.)

Though the distinction between perception and judgment is often clear and intuitive – in part because they can so clearly conflict (as in visual illusions) – we contend that judgment-based alternative explanations for top-down effects have been severely underappreciated in recent work. Fortunately, there are straightforward approaches for teasing them apart.

4.2.1. Case studies. It has been reported that throwing a heavy ball (rather than a light ball) at a target increases estimates of that target's distance (Witt et al. 2004). One interpretation of this result (favored by the original authors) is that the increased throwing effort actually made the target *look* farther away, and that this is why subjects gave greater distance estimates. However, another possibility is that subjects only judged the target to be farther, even without a real change in perception. For example, after having such difficulty reaching the target with their throws, subjects may have simply concluded that the target must have been farther away than it looked.

A follow-up study tested these varying explanations and decided the issue in favor of an effect on judgment rather than perception. Whereas the original study asked for estimates of distance without specifying precisely how subjects should make such estimates, Woods et al. (2009) systematically varied the distance-estimation instructions, contrasting cases (between subjects) asking for reports of how far the target "visually appears" with cases asking for reports of "how far away you feel the object is, taking all nonvisual factors into account" (p. 1113). In this last condition, subjects were especially encouraged to separate perception from judgment: "If you think that the object appears to the eye to be at a different distance than it feels (taking nonvisual factors into account), just base your answer on where you feel the object is." Tellingly, the effect of effort on distance estimation replicated only in the "nonvisual factors" group, and not in the "visually appears" group – suggesting that the original results reflected what subjects *thought* about the distance rather than how the distance truly *looked* to them.

Similarly, it was reported that accuracy in throwing darts at a target affected subsequent size judgments of the target, which was initially assumed to reflect a perceptual change: Less-accurate throwing led to smaller target-size estimates, as if one's performance perceptually resized the target (Wesp et al. 2004; Fig. 2F). However, the same researchers rightly wondered whether this was genuinely an effect on perception or whether these biased size estimates might instead be driven by overt inferences that the target must have been smaller than it looked (perhaps to explain or justify poor throwing performance). To test this alternative, the same research group (Wesp & Gasper 2012) replicated the earlier result – but then ran a follow-up condition in which, before throwing, subjects were told that the darts were faulty and inaccurate. This additional instruction eliminated the correlation between performance and reported size. With a ready-made explanation already in place, subjects no longer needed to "blame the target": Rather than conclude that their poor throwing resulted from a small target, subjects instead attributed their performance to the supposedly faulty darts and thus based their size estimates directly on how the target looked.

Note that this is a perfect example of the *kind* of judgment that can only be described as "routine." Even if other sorts of top-down effects on judgment more richly interact with foundational issues in perception research, blaming a target for one's poor performance is not one of them.

4.2.2. Other susceptible studies. Many alleged top-down effects on perception seem explicable by appeal to these sorts of routine judgments. One especially telling pattern of results is that many of these effects are found even when no visual stimuli are used at all. For example, factors such as value and ease of action have been claimed to affect online distance perception (e.g., Balcetis & Dunning 2010; Witt et al. 2005), but those same factors have been shown to affect the estimated distance of completely unseen (and sometimes merely imagined) locations such as Coney Island (Alter & Balcetis 2011) or one's work office (Wakslak & Kim 2015). Clearly such effects *must* reflect judgment and not perception – yet their resemblance to other cases that are indeed claimed as top-down effects on perception suggests that many such cases could reflect judgmental processes after all.

Other cases seem interpretable along these lines all on their own. For example, another study also demonstrated an effect of dart-throwing performance on size judgments – but found that the effect disappeared when subjects made their throws while hanging onto a rock-climbing wall 12 feet above the ground (Cañal-Bruland et al. 2010). Though this phenomenon was interpreted as an effect of anxiety on action-specific perception, the finding could easily be recast as an effect on judgment instead: Subjects who performed poorly while clinging to the rock-climbing wall had an obvious explanation for performing poorly and so had no need to explain their misses by inflating target-size estimates.

In other cases, the inference to judgment rather than perception can be more straightforward. For example, politically conservative subjects rated darkened images of Barack Obama as more "representative" of him than lightened images, whereas liberal subjects showed the opposite pattern (Caruso et al. 2009), and this effect was interpreted as an effect of partisan attitudes on perceived skin tone. However, it seems more likely that darker photos (or darker skin tones) seemed more negative to subjects, and that conservatives deemed them more representative (and liberals less representative) for that reason – because conservatives think more negatively about Obama than liberals do. (By analogy, we suspect that conservative subjects would also rate a doctored image of Obama with bright red horns on his forehead as more "representative" than an image of Obama with a halo, and that liberals would show the opposite pattern; but clearly such a result would not imply that conservatives literally *see* Obama as having horns!) Other purported top-down effects that seem similarly explicable include effects on visually estimated weight (Doerrfeld et al. 2012), the estimated looming of spiders (Riskind et al. 1995), and the rated anger in African-American or Arab faces (Maner et al. 2005).

4.2.3. A lesson for future research. The distinction between perception and judgment is intuitive and uncontroversial in principle, but it is striking just how few discussions of top-down effects on perception even mention judgmental effects as possible alternative explanations. (For some exceptions see Alter & Balcetis 2011; Lupyan et al. 2010; Witt et al. 2010.) Future work relying on subjective perceptual reports must attempt to disentangle these possibilities. It would of course be preferable for such studies to empirically distinguish perception from judgment – for example, by using performance-based measures in which subjects' success is tied directly to how they perceive the stimuli (such as a visual search task; cf. Scholl & Gao 2013). Or, per the initial case study reviewed above, future work can at least ask the key questions in multiple ways that differentially load on judgment and perception.

At a minimum, given the importance of distinguishing judgment from perception, it seems incumbent on any proposal of a top-down effect to explicitly and prominently address the distinction, even if only rhetorically – because a shift from perception to judgment may dramatically reduce such an effect's potential revolutionary consequences. And at the same time, we note that certain terms may actively obscure this issue and so should be avoided. For example, many papers in this literature advert to effects on "perceptual judgment" (e.g., Meier et al. 2007; Song et al. 2012; Storbeck & Stefanucci 2014), which can only invite confusion about this foundational distinction.

4.3. Pitfall 3: Demand and response bias

Vision experiments occur in a variety of controlled environments (including the laboratory), but any such environment is also inevitably a social environment – which raises the possibility that social biases may intrude on perceptual reports in a more specific way than we saw in Pitfall 2. Whereas *judgments* of various visual qualities are often sincerely held even when they are subject to top-down influence (such that, e.g., inaccurate dart-throwers may truly believe that the target must be smaller than it looks), other sorts of biases may reflect more active modulation of responses by participants – such that this pitfall is conceptually distinct from the previous one. In particular, the social nature of psychology experiments can readily lead to reports (of anything, including percepts) being contaminated by *task demands,* wherein certain features of experiments lead subjects to adjust their responses (either consciously or unconsciously) in accordance with their assumptions about the experiment's purpose (or the experimenters' desires). (For a review of the power and pervasiveness of such effects, see Rosenthal & Rubin 1978.)

Contamination by demand characteristics seems especially likely in experiments involving a single conspicuous manipulation and a single perceptual report. But even more so than with the previous pitfall, it seems especially easy to combat such influences – for example, by asking subjects directly about the experiment and/or by directly manipulating their expectations.

4.3.1. Case studies. Consider the effect of wearing a heavy backpack on slant estimates (Bhalla & Proffitt 1999; Fig. 2D). One possibility is that backpacks make hills look steeper, and that the subjects faithfully reported what they saw. But another explanation is that subjects modified their responses to suit the experimental circumstances, in which a very conspicuous manipulation (a curiously unexplained backpack) was administered before obtaining a single perceptual judgment (regarding the hill's slant).

A recent series of studies shows that the experimental demand of wearing a backpack can completely account for the backpack's effect on slant estimates. When backpack-wearing subjects were given a compelling (but false) cover story to justify the backpack's purpose (to hold heavy monitoring equipment during climbing), the effect of heavy backpacks on slant estimation completely disappeared (Durgin et al. 2009; see also Durgin et al. 2012). With a plausible cover story, subjects had very different expectations about the experiment's purpose (expectations that they articulated explicitly during debriefing), which no longer suggested that the backpack "should" modulate their responses. Similar explanations have subsequently been confirmed for other effects of action on perceptual reports, including effects of aperture "passability" on spatial perception (Firestone & Scholl 2014b) and energy on slant perception (Durgin et al. 2012; Shaffer et al. 2013). For example, no effect of required climbing effort is found *without* a transparent manipulation – such as when subjects estimate the slant of either an (effortful) staircase or an (effort-free) escalator in a between-subjects design (Shaffer & Flint 2011).

Other studies have implicated task demands in very different top-down effects. For example, it has been reported that, when subjects can win a gift card taped to the ground if they throw a beanbag closer to the gift card than their peers do, subjects undershoot the gift card if it is worth $25 but not if it is worth $0 – suggesting (to the original authors) that more desirable objects look closer (Balcetis & Dunning 2010; Fig. 2P). However, in addition to the value of the gift card, the demands of the task differed across these conditions in an especially intuitive way: Whereas subjects may employ various throwing strategies in earnest attempts to win a $25 gift card, they may not try to "win" a $0 gift card (which is a decidedly odd task). For example, subjects who are genuinely trying to win the $25 gift card might undershoot the card if they believed it would be awarded to the closest throw without going over, or if they anticipated that the beanbag would bounce closer to the gift card after its first landing. However, they may not show those biases for the $0 gift card, which wouldn't have been worth any such strategizing. A follow-up study (Durgin et al. 2011a) tested these possibilities directly and found that slightly changing the instructions so that the challenge was to hit the gift card directly (rather than land closest) led subjects to throw the beanbag farther (perhaps because they were no longer worried that it would bounce or that they would be disqualified if they overshot), just as would be expected if differences in strategic throwing (rather than differences in actual perception) explained the initial results.

4.3.2. Other susceptible studies. Perhaps no pitfall is as generally applicable as demand and response bias, especially for studies relying entirely on observer reports. A great many reported top-down effects on perception use very salient manipulations and ask for perceptual judgments that either give subjects ample opportunity to consider the manipulation's purpose or make the "right" answer clear. For example, it has been reported that, when shown a range of yellow-orange discs superimposed on a traffic light's middle bulb, German subjects (for whom that light's color is called *gelb,* or yellow) classified more discs as "yellow" than did Dutch subjects (who call it *oranje,* or orange; Mitterer et al. 2009; Fig. 2Q). Though interpreted as an effect of language on perception – the claim being that the German subjects visually experienced the colored discs as yellower – it seems just as plausible that the subjects were simply following convention, assigning the yellow-orange discs the socially appropriate names for that context.

Many other studies use salient manipulations and measures in a manner similar to the backpacks and hills experiments (Bhalla & Proffitt 1999). For example, similar explanations seem eminently plausible for reported effects of desirability on distance perception (e.g., the estimated distance of feces vs. chocolate; Balcetis & Dunning 2010), of racial identity on faces' perceived lightness (Levin & Banaji 2006), of stereotypes on the identity of weapons and tools (Correll et al. 2015), of tool use on the perceived distance to reachable targets (Witt et al. 2005), of scary music on the interpretation of scary or nonscary ambiguous figures (Prinz & Seidel 2012; Fig. 2H), and of fear of heights on perceived height (Clerkin et al. 2009; Stefanucci & Proffitt 2009).

4.3.3. A lesson for future research. In light of recent findings concerning task demands in studies of top-down effects on perception (especially Durgin et al. 2009), it is no longer possible to provide compelling evidence for a top-down effect on perception without considering the experiment's social context. Yet, so many studies never even mention the possibility of demand-based effects (including several studies mentioned above, e.g., Mitterer et al. 2009; Prinz & Seidel 2012). (For some exceptions, see Levin & Banaji 2006; Schnall et al. 2010; Witt 2011b.) This is especially frustrating because assessing demand effects is often easy and cost-free. In particular, although demand effects can be mitigated by nontransparent manipulations or indirect measures, they can also often be assessed by simply *asking the subjects about the experiment* – for example, during a careful postexperiment debriefing. For example, before the experiment's purpose is revealed, researchers can carefully ask subjects what they thought the experiment was about, what strategies they used, and so forth. These sorts of questions can readily reveal (or help rule out) active demand factors.

Such debriefing was especially helpful, for example, in the case of backpacks and reported slant, wherein many subjects explicitly articulated the experimental hypothesis when asked – and only those subjects showed the backpack effect (Durgin et al. 2009). In this way, we believe Durgin et al.'s 2009 report has effectively set the standard for such experiments: Given the negligible costs and the potential intellectual value of such careful debriefing, we contend that claims of top-down effects (especially in studies using transparent manipulations) can no longer be credible without at least asking about – and reporting – subjects' beliefs about the experiment.

4.4. *Pitfall 4: Low-level differences (and amazing demonstrations!)*

Whereas many studies search for top-down effects on perception by manipulating states of the perceiver (e.g., motivations, action capabilities, or knowledge), many other top-down effects involve manipulations of the stimuli used across experimental conditions. For example, one way to test whether arousal influences spatial perception could be to test a high-arousal group and a low-arousal group on perception of the same stimulus (e.g., a precarious height; Teachman et al. 2008). However, another strategy could be to measure how subjects perceive the distance of arousing versus nonarousing stimuli (e.g., live tarantulas vs. plush toys; Harber et al. 2011). Though both approaches have strengths and weaknesses, one difficulty in manipulating stimuli across experimental conditions is the possibility that the intended top-down manipulation (e.g., the evoked arousal) is confounded with changes in the low-level visual features of the stimuli (e.g., as live tarantulas might differ in size, color, and motion from plush toys) – and that those low-level differences might actually be responsible for perceptual differences across conditions.

We have suggested (and will continue to suggest) that many of the pitfalls we discuss here have been largely neglected by the literature on top-down effects, but this pitfall is an exception: Studies that manipulate stimuli often do acknowledge the possibility of low-level differences (and on occasion actively attempt to control for them). Nevertheless, we contend that such low-level differences are even more pervasive and problematic than has been realized, and that simple experimental designs can reveal when such differences are responsible for apparent top-down effects.

4.4.1. Case studies. One especially compelling and currently influential top-down effect on perception is a report that Black (i.e., African-American) faces look darker than White (i.e., Caucasian) faces, even when matched for mean luminance, as in Figure 3A (Levin & Banaji 2006). This finding is today widely regarded as one of the strongest counterexamples to modularity (e.g., Collins & Olson 2014; Macpherson 2012; Vetter & Newen 2014) – no doubt because, in addition to the careful experiments reported in the paper, the difference in lightness is clearly apparent upon looking at the stimuli. In other words, this top-down effect works as a "demonstration" as well as an experiment.

That last point is worth emphasizing given the prevalence of "demos" in vision science. In our field, experimental data about what we see are routinely accompanied by such demonstrations – in which interested observers can experience the relevant phenomena for themselves, often in dramatic fashion. For example, no experiments are *needed* to convince us of the existence of phenomena such as motion-induced blindness, apparent motion, or the lightness illusions depicted in Figure 1. Of course, that is not to say that demos are necessary for progress in vision science; most experiments surely get by without them. But effective demos can provide especially compelling evidence, and they may often directly rule out the kinds of worries we expressed in discussing the previous two pitfalls (i.e., task demands and postperceptual judgments).

In this context, the demonstration of race-based lightness distortions (Levin & Banaji 2006) is exceptional, insofar as it is one of the only such demos in this large literature. Indeed, it strikes us as an awkward fact that so few such effects can actually be experienced for oneself. For example, the possibility that valuable items look closer is testable not only in a laboratory (e.g., Balcetis & Dunning 2010) but also from the comfort of home: Right now you can place a $20 bill next to a $1 bill and see for yourself whether there is a perceptual difference. Similarly, knowledge of an object's typical color (e.g., that bananas are yellow) reportedly influences that object's perceived color, such that a grayscale image of a banana is judged to be more than 20% yellow (Hansen et al. 2006; Olkkonen et al. 2008); however, if you look now at a grayscale image of a banana (Fig. 2K), we predict that you will not experience this effect for yourself – even though the reported effect magnitudes far exceed established discrimination thresholds (e.g., Hansen et al. 2008; Krauskopf & Gegenfurtner 1992). (You may notice that many of the top-down effects in Figure 2 are caricatured, e.g., with *actual* luminance differences for positive vs. negative words and smiling vs. frowning faces. This is because when the effects weren't caricatured in this way, readers could not understand the claims – because they could not experience the effect!)

All of this makes the reported lightness difference much more compelling: As you may experience in Figure 3A, the Black face truly looks darker than the luminance-matched White face. But is this a *top-down* effect on perception? Though the face stimuli were matched for mean luminance, there are of course many visual cues to lightness that are independent of mean luminance. For example, in many lightness illusions, two regions of equal luminance nevertheless appear to have different lightnesses because of depicted patterns of illumination and shadow (as in Fig. 1). Indeed, a close examination of the face stimuli in Figure 3A suggests that the Black face seems to be under illumination, whereas the White face doesn't look particularly illuminated or shiny – a difference that has long been known to influence perceived lightness (Adelson 2000; Gilchrist & Jacobsen 1984). And the Black face has a darker jawline, whereas the White face has darker eyes. Of course, there must exist *some* low-level differences between the images, because otherwise they would be identical; nevertheless, the question remains whether such lower-level visual factors are responsible for the effect, rather than the meaning or category (here, race) that is correlated with that low-level difference.

To test whether one or more such low-level differences – rather than race, per se – explain the difference in perceived lightness, we replicated this study with blurred versions of the face stimuli, so as to eliminate race information while preserving many low-level differences in the images (including the match in average luminance and contrast) – as in Figure 3B (Firestone & Scholl 2015a). After blurring, the vast majority of observers asserted that the two faces actually had the same race (or were even the same person). However, even those observers who asserted that the faces had the same race nevertheless judged the blurry image derived from the Black face to be darker than the blurry image derived from the White face. This result

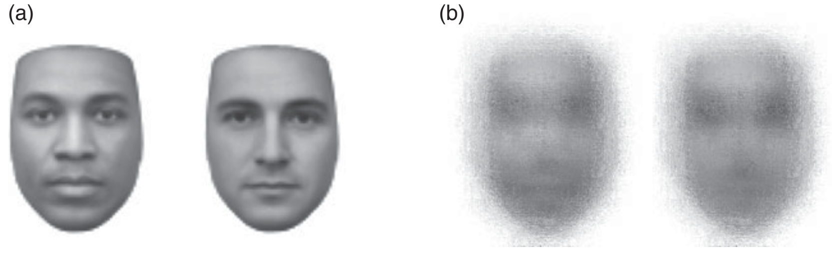


Figure 3. (**A**) The face stimuli (matched in mean luminance) from Experiment 1 of Levin and Banaji (2006). (**B**) The blurred versions of the same stimuli used in Firestone and Scholl (2015a), which preserved the match in mean luminance but obscured the race information.

effectively shows how the lightness difference can derive from low-level visual features – which, critically, are present in the original images – without any contribution from perceived race. (And note that such results are unlikely to reflect *unconscious* race categorization; it would be a distinctively odd implicit race judgment that could influence explicit lightness judgments *but not explicit race judgments*.) And although the original effect (with unblurred faces) could of course still be explained entirely by race (rather than by the lower-level differences now shown to affect perceived lightness), it is clear that further experiments would be needed to show this – and so we conclude that the initial demonstration of Levin and Banaji (2006) provides no evidence for a top-down effect on perception.[3]

Many other effects that initially seemed to reflect high-level factors have been shown to reflect lower-level visual differences across conditions. Consider, for example, reports that categorical differences facilitate visual search with line drawings of animals and artifacts – e.g., with faster and more efficient searches for animals among artifacts, and vice versa (Levin et al. 2001). This result initially appeared to be a high-level effect on a fairly low-level perceptual process, given that efficient visual search is typically considered "preattentive." However, on closer investigation (and to their immense credit), the same researchers discovered systematic low-level differences in their stimuli – wherein the animals (e.g., snakes, fish) had more curvature than the artifacts (e.g., chairs, briefcases) – which sufficiently explained the search benefits (as revealed by follow-up experiments directly exploring curvature).

4.4.2. Other susceptible studies. The possibility of such low-level confounds is a potential issue, almost by definition, with any top-down effect that varies stimuli across conditions. For example, size contrast is reportedly enhanced when the inducing images are conceptually similar to the target image – such that a central dog image looks smaller when surrounded by images of larger dogs versus images of larger shoes (Coren & Enns 1993). However, dogs are also more geometrically similar to each other than they are to shoes (for example, the shoe images were shorter, wider, and differently oriented than the dog images; Fig. 2L), and size contrast may instead be influenced by such geometric similarity. (Coren and Enns are quite sensitive to this concern, but their follow-up experiments still involve important geometric differences of this sort.)

Other investigations of top-down effects on size contrast also manipulate low-level properties, for example contrasting natural scenes with different distributions of color and complexity (e.g., van Ulzen et al. 2008). Or, in a very different case, studies of how fear may affect spatial perception often involve stimuli with very different properties (e.g., a live tarantula vs. a plush toy; Harber et al. 2011), or even the same stimulus viewed from very different perspectives (e.g., a precarious balcony viewed from above or below; Stefanucci & Proffitt 2009).

4.4.3. A lesson for future research. Manipulating the actual stimuli or viewing circumstances across experimental conditions is a perfectly viable methodological choice, but it adds a heavy burden to avoid low-level differences between the stimuli. Critically, this burden can be met in at least two ways. One possibility is to preserve the high-level factor while eliminating the low-level factor. (In other contexts looking at fearful stimuli, for example, images of spiders have been contrasted not with plush toys, but with "scrambled" spider images, or even images of the same line segments rearranged into a flower; e.g., New & German 2015.) Another possibility – as in our study of race categories and lightness (Firestone & Scholl 2015a) – is to preserve the low-level factor while eliminating the high-level factor. For top-down effects, this latter strategy is often more practical because it involves positively replicating the relevant effect; in contrast, the former strategy may require a null effect (which raises familiar concerns about statistical power, etc.). In either case, however, such strategies show how this pitfall is eminently testable.

4.5. Pitfall 5: Peripheral attentional effects

We have been arguing that there are no top-down effects of cognition on perception, in the strong and revolutionary sense wherein such effects violate informational encapsulation or cognitive impenetrability and so threaten the view of the visual system as a functionally independent (modular) part of the mind. However, we have also noted some other senses of top-down effects that carry no such implications (see sect. 2). Chief among these is the notion of changing what we see by changing the input to perception, as when we close (or move) our eyes based on our desires (see sect. 2.1).

Other ways of changing the input to perception, however, are more subtle. Perhaps most prominently, shifting patterns of *attention* can change what we see. Selective attention is obviously closely linked to perception – often serving as a gateway to conscious awareness in the first place, such that we may completely fail to see what we do not attend to (as in inattentional blindness; e.g., Most et al. 2005b; Ward & Scholl 2015). Moreover, attention – which is often likened to a "spotlight" or "zoom lens" (see Cave & Bichot 1999; though cf. Scholl 2001) – can sometimes literally highlight or enhance attended objects, making them appear (relative to unattended objects) clearer (Carrasco et al. 2004) and more finely detailed (Gobell & Carrasco 2005).

Attentional phenomena relate to top-down effects simply because attention is at least partly under intentional control – insofar as we can often choose to pay attention to one object, event, feature, or region rather than another. When that happens – say, if we attend to a specific flower and it looks clearer or more detailed – should that not then count as our intentions changing what we see?

In many such cases, changing what we see by selectively attending to a different object or feature (e.g., to people passing a basketball rather than to a dancing gorilla, or to black shapes rather than white shapes; Most et al. 2005b; Simons & Chabris 1999) seems importantly similar to changing what we see by moving our eyes (or turning the lights off). In both cases, we are changing the input to mechanisms of visual perception, which may then still operate inflexibly given that input. A critical commonality, perhaps, is that the influence of attention (or eye movements) in such cases is completely independent of your *reason* for attending that way. Having the lights turned off has the same effect on visual perception regardless of why the lights are off, including whether you turned them off intentionally or accidentally; in both cases it's the change in the light doing the work, not the antecedent intention. And in similar fashion, attention may enhance what you see regardless of the reasons that led you to deploy attention in that way, and even whether you attended voluntarily or through involuntary attentional capture; in both cases, it's the change in attention doing the work, not the antecedent intention. Put differently, such attentional (or light-turning-off) effects may be occasioned by a relevant intention or belief, but they are not sensitive to the *content* of that intention or belief.

Moreover, such attentional effects are already part of the "orthodoxy" in vision science, which currently studies and models such attentional effects and readily accepts that shifts in attention can affect what we see. By contrast, a primary factor that makes other top-down effects (e.g., effects of morality, hunger, language, etc. on perception) potentially revolutionary in the present context is precisely that they are not part of this traditional understanding of visual perception.

Of course, not all attentional effects must be so peripheral in nature. In other contexts, attention may interact in rich and nuanced ways with unconscious visual representations to effectively mold and choose a "winning" percept – changing the content of perception rather than merely influencing what we focus on. (For an elaboration of how such attentional dynamics may interact with issues of modularity, see Clark 2013). However, our contention in this pitfall is that the merely *peripheral* sorts of attention – involving simple changes in which locations, features, or objects we focus on – can account for a wide variety of apparent top-down effects on perception. As a result, we focus on such peripheral forms of attention in the rest of this section, while not denying that attention can also interact with perception in much richer ways as well.

In light of such considerations, it seems especially important to determine for any alleged top-down state (e.g., an intention, emotion, or desire) whether that state is influencing what we see *directly* (in which case it may undermine the view that perception is a functionally independent module) or whether it is (merely) doing so indirectly by changing how we attend to a stimulus in relatively peripheral ways – in which case it may simply change the input to visual processing but not how that processing operates.

4.5.1. Case studies. Attention has a curious status in the long-running debate about top-down effects. On the one hand, perhaps based on its prominence in previous discussions (especially Pylyshyn 1999), the kinds of thoughts noted in the previous section are almost always recognized and accepted in most modern discussions of top-down effects – including recent discussions reaching very different conclusions than our own (e.g., two of the most recent literature reviews of top-down effects, which concluded that top-down effects have been conclusively established many times over; Collins & Olson 2014; Vetter & Newen 2014). Nevertheless, these recent discussions happily agree that, to be compelling, top-down effects must not merely operate through attention. (Collins and Olson draw their conclusions based largely on contemporary top-down effects – including Levin and Banaji 2006 – that, unlike earlier results, "cannot be easily attributed to the influences of attention," p. 843. And Vetter and Newen even define the question in terms of top-down effects obtaining "while the state of the sensory organs (in terms of spatial attention and sensory input) is held constant," p. 64.)

On the other hand, given this prominence in theoretical discussions of top-down effects, it is curious that attention is almost never empirically explored in this literature – curious especially given that such effects are often straightforwardly testable. In particular, for a possible influence of intention on perception (for instance), it is almost always possible to separate intention from attention – most directly by holding one constant while varying the other. For example, to factor attention out, one can impose an attentional load, often by means of a secondary task (which is common in many other contexts, e.g., in exploring scene perception or feature binding without attention; Bouvier & Treisman 2010; Cohen et al. 2011). Or, in a complementary way, one can assess attention's role directly by actively manipulating the locus of attention while holding intention constant. This is what was done in one of the only relevant empirical case studies of attention and top-down effects.

Perhaps the most intuitively compelling evidence for intentional effects on perception comes from studies of ambiguous figures. For example, inspection of a Necker cube (Fig. 2R) reveals that one can voluntarily switch which interpretation is seen (in particular, which face of the cube seems to be in front), and such switching is also possible with many other such figures (such as the famous duck–rabbit figure). Several early defenders of top-down influences on perception essentially took such intuitions at face value and rejected the cognitive impenetrability of visual perception from the get-go. For example, Churchland (1988) argued that one controls which interpretations of these ambiguous images one sees by "changing one's assumptions about the nature of the object," and thus concluded that "at least some aspects of visual processing, evidently, are quite easily controlled by the higher cognitive centers" (p. 172).

However, several later studies showed that such voluntary switches from one interpretation to another are occasioned by exactly the sorts of processes that uncontroversially do *not* demonstrate top-down penetration of perception. For example, switches in the Necker cube's interpretation are driven by changes in which corner of the cube is attended (Peterson & Gibson 1991; Toppino 2003), and the same has been found for other ambiguous figures (for a review, see Long & Toppino 2004). (Such effects are driven by the fact that attended surfaces tend to be seen as closer.) In other words, though one may indeed be "changing one's assumptions" when the figure switches, that is not actually triggering the switches. Instead, the mechanism is that different image regions are selectively processed over others, because such regions are attended differently in relatively peripheral ways.

Evidence has recently emerged pointing to a similar explanation for certain effects of action on spatial perception. It has been reported that success in a golfing task inflates perception of the golf hole's size (Witt et al. 2008), and more generally that successful performance makes objects look closer and larger (Witt 2011a). However, follow-up studies suggest that the mere deployment of attention may be the mediator of these effects. For example, diverting attention away from the hole at the time the golf ball is struck (by occluding the hole, or by making subjects putt around the blades of a moving windmill) eliminates the effect of performance on judged golf-hole size (Cañal-Bruland et al. 2011), and the presence of the hole-resizing effect is associated with a shift in the location of subjects' attentional focus (e.g., toward the club vs. toward the hole; Gray & Cañal-Bruland, 2015; see also Gray et al. 2014 for a similar result in the context of a different action-specific top-down effect). Moreover, the well-documented effects of action-planning on spatial judgments (e.g., Kirsch & Kunde 2013a; 2013b) also arise from the deployment of visual attention alone, even in the absence of motor planning (Kirsch 2015). That visual attention may be both necessary and sufficient for such effects suggests that apparent effects of action on perception may reduce to more routine and well-known interactions between attention and perception.

4.5.2. Other susceptible studies. There have been fewer case studies of this sort of peripheral attention and its role in top-down effects, and as a result this pitfall is not on the sort of firm empirical footing enjoyed by the other pitfalls presented here. Nevertheless, it seems that such explanations could apply broadly. Most immediately, many recently reported top-down effects on perception use ambiguous figures but do not rule out relevant attention mechanisms. For example, it has been reported that scary music biased the interpretation of ambiguous figures (such as an alligator/squirrel figure; Fig. 2H) toward their scarier interpretation (Prinz & Seidel 2012), and that subjects who are rewarded every time a certain stimulus is shown (e.g., the number 13) report seeing whichever interpretation of an ambiguous stimulus (e.g., a B/13 figure) is associated with that reward (Balcetis & Dunning 2006). However, such studies fail to measure attention, or even (in the case of Prinz & Seidel) to mention it as a possibility.

The effects of attention on appearance pose an even broader challenge. For example, findings suggesting that attended items can be seen as larger (Anton-Erxleben et al. 2007) immediately challenge the interpretation of nearly every alleged top-down effect on size perception, including reported effects of throwing performance on perceived target size (e.g., Cañal-Bruland et al. 2010; Fig. 2F), of balance on the perceived size of a walkable beam (Geuss et al. 2010), of hitting ability on the perceived size of a baseball (Gray 2013; Witt & Proffitt 2005), of athletes' social stature on their perceived physical stature (Masters et al. 2010), and even of sex primes on the perceived size of women's breasts (den Daas et al. 2013). In this last case, for example, if sex-primed subjects simply attended more to the images of women's breasts, then this could explain why they (reportedly) appeared larger. And this is to say nothing of attentional effects on other visual properties – such as the fact that voluntary attention perceptually darkens transparent surfaces (Tse 2005), which could explain the increase in perceived transparency among thirsty observers (Changizi & Hall 2001; Fig. 2B) if nonthirsty observers simply paid more attention during the task (being less distracted by their thirst).

4.5.3. A lesson for future research. Attention is a rich and fascinating mental phenomenon that in many contexts interacts in deep and subtle ways with foundational issues in perception research. However, there also exist more peripheral sorts of attention that amount to little more than focusing more or less intently on certain locations (or objects, or features) in the visual field. And because such peripheral forms of attention are ubiquitous and active during almost every waking moment, future work must rule out peripheral forms of attention as mediators of top-down effects in order to have any necessary implications for the cognitive impenetrability of visual perception.

Given how straightforward such tests are in principle (per sect. 4.5.1), studies of attention could play a great role in advancing this debate – either for or against the possibility of truly revolutionary top-down effects. Some top-down effects might be observed even when attention is held constant or otherwise occupied, and this would go a long way toward establishing them as counterexamples to the modularity of perception. Or, it could be shown that the deployment of visual attention alone is insufficient to produce similar effects. For example, if it were shown that attending to a semitransparent surface does not make it look more opaque, this could rule out the possibility that attention drove thirst-based influences on perceived transparency (Changizi & Hall 2001). Similarly, if moving one's attention from left to right fails to bias apparent motion in that direction, then such attentional anticipation may not underlie alleged effects of language on perceived motion (as in Meteyard et al. 2007; Tse & Cavanagh 2000). Such studies would be especially welcome in this literature, given the seemingly dramatic disparity between how often attention is theoretically recognized as relevant to this debate and how seldom it is empirically studied or ruled out.

4.6. Pitfall 6: Memory and recognition

Top-down effects on perception are meant to be effects on what we *see*, but many such studies instead report effects on how we *recognize* various stimuli. For example, it has been reported that assigning linguistic labels to simple shapes improves reaction time in visual search and other recognition tasks (Lupyan & Spivey 2008; Lupyan et al. 2010; Fig. 2M), and that, when briefly presented, morally relevant words are easier to identify than morally irrelevant words (the "moral pop-out effect" as in Fig. 2C; Gantman & Van Bavel 2014; see also Radel & Clément-Guillotin 2012). Such reports often invoke the revolutionary language of cognitive penetrability (Lupyan et al. 2010) or claim effects on "basic awareness" or "early" perceptual processing (Gantman & Van Bavel 2014; Radel & Clément-Guillotin 2012). However, by its nature, recognition necessarily involves not only visual processing per se, but also memory: To recognize something, the mind must determine whether a given visual stimulus matches some stored representation in memory. For this reason, any top-down improvement in visual recognition could reflect a "front-end" effect on visual processing itself (in which case such effects would indeed have the advertised revolutionary consequences), or instead a "back-end" effect on memory access (in which case they would not, if only because many top-down effects on memory are undisputed and even pedestrian).

Of course, other sorts of top-down effects on memory may interact with perception in richer and more intimate ways. For example, simply holding an item in visual working memory may speed awareness of that object – as when a stimulus that matches a target held in memory is quicker to escape continuous flash suppression (Pan et al. 2014) or motion-induced blindness (Chen & Scholl 2013). But our contention is that even those phenomena of "back-end" memory with no intrinsic connection to seeing or cognitive penetrability – such as spreading activation in semantic memory – can explain many alleged top-down effects on perception. And often, this contrast between front-end perception and mere back-end memory is directly testable.

4.6.1. Case studies. Consider again the "moral pop-out effect" – the report that morally relevant words are easier to see than morally irrelevant words, supposedly because moral stimuli are "privileged" in the mind (Gantman & Van Bavel 2014). Subjects were shown very briefly (40–60 ms) presented words and nonwords one at a time and had to decide whether each presented stimulus was a word or a nonword (see Fig. 4A). Some words were morally relevant (e.g., "illegal"), and some were morally irrelevant (e.g., "limited"). Subjects more accurately identified morally relevant words than morally irrelevant words. However, by virtue of being related to morality, the morally relevant words were also related to *each other* (e.g., the moral words included not only "illegal" but also "law," "justice," "crime," "convict," "guilty," and "jail"), whereas the non-moral words were not related to anything in particular (including, in addition to "limited," words such as "rule," "exchange," "steel," "confuse," "tired," and "house"). In that case, it could be that the moral words simply primed each other and were easier to recognize for that reason rather than because of anything special about morality.

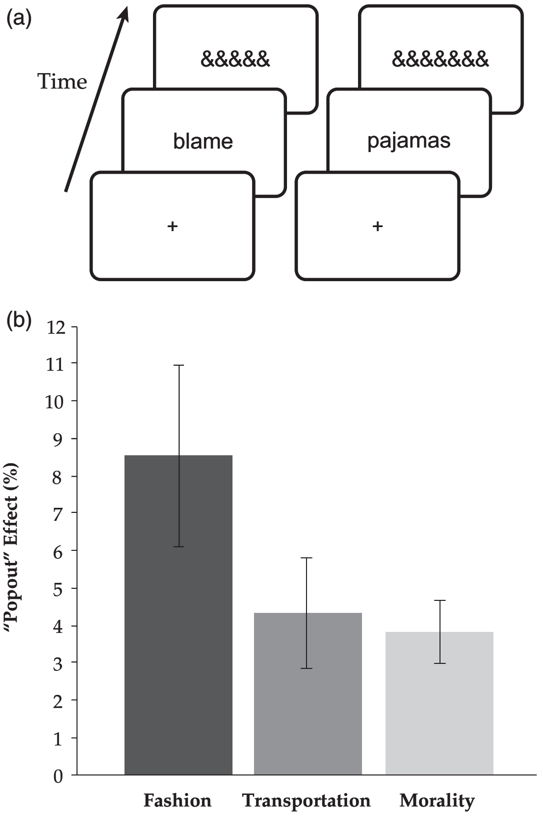


Figure 4. (A) Experimental design for detecting "enhanced visual awareness" of various word categories. After fixation, a word or nonword appeared for 50 ms and then was masked by ampersands for 200 ms. (B) Results comparing "popout effects" for fashion, transportation, and morality (from Firestone & Scholl 2015b).

Crucially, such semantic priming would not be a top-down effect on *perception*. For example, in more traditional lexical decision tasks (e.g., Meyer & Schvaneveldt 1971), in which visual recognition of a word (e.g., "nurse") is speedier and more accurate when that word is preceded by a related word (e.g., "doctor") than by an unrelated word (e.g., "butter"), many follow-up experiments and modeling approaches have shown that this improvement in recognition occurs not because of any boost to visual processing per se, but rather because the relevant memory representations are easier to retrieve when evaluating whether a visual stimulus is familiar (e.g., because "doctor" activates semantically related lexical representations in memory, including "nurse"; Collins & Loftus 1975; Masson & Borowsky 1998; Norris 1995). Thus, just as "doctor" primes semantically related words such as "nurse," words such as "illegal" may have primed related words such as "justice" (whereas words such as "limited" would not have primed unrelated words such as "exchange").

One unique prediction of this alternative account is that, if the results are driven simply by semantic relatedness and its effect on memory, then *any* category should show a similar "pop-out effect" in similar circumstances – including completely arbitrary categories without any special importance in the mind. To test this possibility, we replicated the "moral pop-out effect" (Firestone & Scholl 2015b), but instead of using words related to morality (e.g., "hero," "virtue"), we used words related to clothing (e.g., "robe," "slippers"). The effect replicated even with this trivial, arbitrary category: Fashion-related words were more accurately identified than non-fashion-related words (see Fig. 4B). (A second experiment replicated the phenomenon again, with transportation-related words such as "car" and "passenger.") These results suggest that relatedness is the key factor in such effects, and thus that memory, not perception, improves detection of words related to morality. In particular, the work done by moral words in such effects (i.e., increased activation in memory) may be complete before any subsequent stimuli are ever presented – just as the spreading activation from "doctor" to "nurse" is complete before "nurse" is ever presented.

Similar investigations have implicated memory in other top-down effects. For example, labeling simple shapes reportedly improves visual detection of those shapes (Lupyan & Spivey 2008). When a certain blocky symbol (⊓⊔) appeared as an oddball item in a search array populated by mirror images of that symbol (⊔⊓), subjects who were told to think of the symbols as resembling the numbers 2 and 5 were faster at finding a ⊓⊔ among ⊔⊓s than were subjects who were given no such special instruction. However, it is again possible that such "labeling" doesn't actually improve visual processing per se but instead makes retrieval of the relevant memory representation easier or more efficient. Indeed, in similar contexts, the prevention of such labeling by verbal shadowing impairs subjects' ability to notice changes to objects' identities (e.g., a Coke bottle changing into a cardboard box) but not to their spatial configuration (e.g., a Coke bottle moving from one location in the scene to another), suggesting that the work done by such labels is to enhance contentful memories rather than visual processing itself (Simons 1996).

In a follow-up, Klemfuss et al. (2012) reasoned that if memory – rather than vision – was responsible for improved detection of a ⊓⊔ among ⊔⊓s, then removing or reducing the task's memory component would eliminate the labeling advantage. To test this account, subjects completed the same task as before, except this time a cue image of the search target was present on-screen during the entire search trial, such that subjects could simply evaluate whether the (still-visible) cue image was also in the search array, rather than whether the search array contained an image held in memory. Under these conditions, the label advantage disappeared, suggesting that memory was the culprit all along. (For another case study of perception vs. memory – in the context of action capabilities and spatial perception – see Cooper et al. 2012.)

4.6.2. Other susceptible studies. Many top-down effects involve visual recognition and so may be susceptible to this pitfall. For example, under continuous flash suppression in a binocular-rivalry paradigm, hearing a suppressed image's name (e.g., the word *kangaroo* when a kangaroo image was suppressed) increased subjects' sensitivity to the presence of the object (Lupyan & Ward 2013) – which they described in terms of the cognitive penetrability of perception. But this too seems readily explicable as an entirely "back-end " memory effect: Hearing the name of the suppressed stimulus activates stored representations of that stimulus in memory (including its brute visual appearance; Léger & Chauvet 2015; Yee et al. 2012), making the degraded information that reaches the subject easier to recognize – whereas, without the label, subjects may have seen faint traces of an image but might have been unable to recognize it as an object (which is what they were asked to report). (Note that this example might thus be explained in terms of spreading activation in memory representations, even with no semantic priming per se.) In another example, hungry subjects more accurately identified briefly presented food-related words relative to subjects who were not hungry (Radel & Clément-Guillotin 2012); but if hungry subjects were simply *thinking about food* more than nonhungry subjects were, then it is no surprise that they better recognized words related to what they were thinking about, having activated the relevant representations in memory even before a stimulus was presented.

4.6.3. A lesson for future research. Given that visual recognition involves both perception and memory as essential but separable parts, it is incumbent on reports of top-down effects on recognition to carefully distinguish between perception and memory, in part because effects on back-end memory have no implications for the nature of perception. (If they did, then the mere existence of semantic priming would have conclusively demonstrated the cognitive penetrability of perception back in the 1970s, rendering the recent bloom of such studies unnecessary.) Yet, it is striking how many recent studies of recognition (including nearly all of those mentioned in sect. 4.6) do not even acknowledge that such an interpretation is important or interesting. (Lupyan & Ward [2013] attempted to guard against semantic priming by claiming that semantic information is not extracted during continuous flash suppression, but recent work now demonstrates that this is not the case; e.g., Sklar et al. 2012. Moreover, Lupyan and Ward worry about semantic priming, but they do not acknowledge the possibility of spreading activation among nonsemantic properties such as visual appearance; see Léger & Chauvet 2015.) Even just highlighting this distinction can help, for example by making salient relevant properties such as relatedness (as in Firestone & Scholl 2015b) when they would be obscure in a purely perceptual context (as in Gantman & Van Bavel 2014). And beyond the necessity of highlighting this distinction, the foregoing case studies also make clear that this is not a vague theoretical or definitional objection but rather a straightforward empirical issue.

5. Discussion and conclusion

There may be no more foundational distinction in cognitive science than that between seeing and thinking. How deep does this distinction run? We have argued that there is a "joint" between perception and cognition to be "carved" by cognitive science, and that the nature of this joint is such that perception proceeds without any direct, unmediated influence from cognition.

Why have so many other scholars thought otherwise? Though many alternative conceptions of what perception is and how it works have deep theoretical foundations and motivations, we suspect that the primary fuel for such alternative conceptions is simply the presence of so many empirical reports of top-down influences on perception – especially in the tidal wave of such effects appearing over the last two decades. (For example, the median publication year of the many top-down reports cited in this paper – which includes New-Look-era reports – is only 2010.) When so many extraperceptual states (e.g., beliefs, desires, emotions, action capabilities, linguistic representations) appear to influence so many visual properties (e.g., color, lightness, distance, size), one cannot help feeling that perception is thoroughly saturated with cognition. And even if one occasionally notices a methodological flaw in one study and a different flaw in another study, it seems unlikely that each top-down effect can be deflated in a different way; instead, the most parsimonious explanation can seem to be that these many studies collectively demonstrate at least *some* real top-down effects of cognition on perception.

However, we have now seen that only a small handful of pitfalls can deflate many reported top-down effects on perception. Indeed, merely considering our six pitfalls – uniquely disconfirmatory predictions, judgment, task demands, low-level differences, peripheral attention, and memory – we have covered at least nine times that many empirical reports of top-down effects (and of course that includes only those top-down effects we had space to discuss, out of a much larger pool).

5.1. Vague theoretical objections and Australian stepbrothers

This is, of course, not the first discussion of potential alternative explanations for top-down effects. Indeed, from the beginning, New Look proponents faced similar criticisms. For example, Bruner and Goodman (1947) note in their original paper on coin-size perception that critics tended to dismiss those findings "by invoking various *dei ex machina*" (p. 33) as alternatives. However, Bruner and Goodman waved these criticisms off: "Like the vengeful and unannounced stepbrother from Australia in the poorer murder mysteries, they turn up at the crucial juncture to do the dirty work.… To shift attention away from [perception] by invoking poorly understood intervening variables does little service" (p. 33). We think this was a perfectly reasonable response: Vague criticisms are cheap and too easy to generate, and it is not a researcher's responsibility to address every far-flung alternative explanation dreamed up off the cuff by anyone who doesn't like some finding. This, however, is where our approach differs sharply and categorically from Bruner and Goodman's would-be Australian stepbrothers. Our six pitfalls are not "poorly understood intervening variables": On the contrary, for each alternative explanation we have offered here, we reviewed multiple case studies suggesting not only that it *could* matter (in principle), but also that it *actually does* matter (in practice) – and applies to many of the most prominent reported top-down effects on perception.

5.2. A checklist for future work

It is our view that no alleged top-down effect of cognition on perception has so far successfully met the challenges collectively embodied by the pitfalls we have reviewed. (No doubt our commentators will educate us otherwise.) Moreover, in the vast majority of cases, it's not that the relevant studies attempted to address these pitfalls but failed; rather, it's that they seem never to have considered most of the pitfalls in the first place. To make progress on this foundational question about the nature of perception, we think future work must take these pitfalls to heart. To this end, we propose that such studies should consider them as a checklist of sorts, wherein each item could be tested (or at least considered) before concluding that the relevant results constitute a top-down effect on perception:

1. *Uniquely disconfirmatory predictions*: Ensure that an effect not only appears where it should, but also that it disappears when it should – for example in situations characterized by an "El Greco fallacy."
2. *Perception versus judgment*: Disentangle postperceptual judgment from actual online perception – for example by using performance-based measures or brief presentations.
3. *Demand and response bias*: Mask the purpose of otherwise-obvious manipulations and measures, and always collect and report subjects' impressions of the experiment's purpose.
4. *Low-level differences (and amazing demonstrations)*: Rule out explanations involving lower-level visual features – for example by careful matching, by directly testing those features without the relevant higher-level factor, or by manipulating states of the observer rather than the stimuli. (And always strive for compelling "demos" of perceptual effects in addition to statistically significant results.)
5. *Peripheral attentional effects*: Either by measuring patterns of attention directly or by imposing an attentional load to attenuate such influences, examine whether higher-level states directly influence lower-level visual processes, or if instead the effect is due to simple changes in which locations, features, or objects are focused on.
6. *Memory and recognition*: When studying top-down influences on recognition, always distinguish "front-end" perception from "back-end" memory, for example by directly varying reliance on memory or actively testing irrelevant categories of stimuli.

Of course, it may not be feasible for every study of top-down effects to conclusively rule out each of these pitfalls. However, such a checklist can be usefully employed simply by taking care to explicitly discuss (or at least mention!) each potential alternative explanation, if only to clarify which alternatives are already ruled out and which remain live possibilities. Doing so would be useful both to opponents of top-down effects (by effectively organizing the possible lines of response) and to their proponents too (by effectively distancing their work from deflationary alternatives). After all, proponents of top-down effects on perception will *want* their effects not to be explained by these pitfalls: If it turns out, for example, that reported effects of desires on perceived size are explained simply by increased attention to the desired object, then such an effect will go from being a revolutionary discovery about the nature of perception to being simply a demonstration that people pay attention to objects they like.

The possibility of top-down effects on perception is tremendously exciting, and it has the potential to ignite a revolution in our understanding of how we see and of how perception is connected to the rest of the mind. Accordingly, though, the bar for a suitably compelling top-down effect should be high. Until this high bar is met, it will remain eminently plausible that there are no top-down effects of cognition on perception.

ACKNOWLEDGMENT
For helpful conversation or comments on earlier drafts, we thank Emily Balcetis, David Bitter, Ned Block, Andy Clark, Frank Durgin, Ana Gantman, Alan Gilchrist, Dan Levin, Gary Lupyan, Dan Simons, Jay Van Bavel, Emily Ward, and the members of the Yale Perception and Cognition Laboratory.

NOTES

1. In practice, researchers have sometimes used these terms in varyingly constrained ways. For example, some discussions of cognitive penetrability (e.g., Pylyshyn 1999; cf. Stokes 2013) emphasize the possibility that a top-down effect may operate via a nonarbitrary "rational connection" more than do discussions of modularity or encapsulation (e.g., Fodor 1983). Here we intend a broad reading of these terms, because they all invoke types of direct top-down effects from higher-level cognitive states on what we see.

2. The same might be said of the oft-cited phenomenon that we see a rainbow as having discrete color bands even though the wavelengths of its light progress continuously through the visible spectrum. Many discussions of this phenomenon of "categorical perception" take pains to note that the effect is preserved even in laboratory studies using psychophysically balanced color spaces, and some assert that, "This is due to our higher-order conceptual representations (in this case, color category labels) shaping the output of color perception" (Collins & Olson 2014, p. 844). But if color category labels truly shape perception in this way, then they should have first shaped the color spaces themselves, leaving no work left for such labels to do upon presentation of the laboratory stimuli using those color spaces.

3. Levin and Banaji (2006) report several other experiments with similar conclusions. None of these other experiments involve subjectively appreciable demonstrations, however, and (in part for this reason) they fall prey to several other pitfalls – including task demands (as in Pitfall 3, because the observers were explicitly told that the study was about "how people perceive the shading of faces of different races"; p. 504) and the El Greco fallacy (as in Pitfall 1, because they observed the effects even when reporting the lightness of a face *using a copy of that exact face*).

Open Peer Commentary

Task demand not so damning: Improved techniques that mitigate demand in studies that support top-down effects

doi:10.1017/S0140525X15002538, e230

Emily Balcetis[a] and Shana Cole[b]

[a]*Department of Psychology, New York University, New York, NY 10003;*
[b]*Psychology Department, Rutgers University, Piscataway, NJ 08854.*
emilybalcetis@nyu.edu **shana.cole@rutgers.edu**
http://www.psych.nyu.edu/balcetis/
http://psych.rutgers.edu/faculty-list-and-links/faculty-profiles-a-contacts/489-shana-cole

Abstract: Firestone & Scholl's (F&S's) techniques to combat task demand by manipulating expectations and offering alternative cover stories are fundamentally flawed because they introduce new forms of demand. We review five superior techniques to mitigate demand used in confirmatory studies of top-down effects. We encourage researchers to apply the same standards when evaluating evidence on both sides of the debate.

Firestone & Scholl (F&S) discuss task demand in study design. Although we agree that researchers should take care to address demand, we fundamentally disagree with F&S's claim that findings of top-down effects on perception can be explained by task demand. They have asymmetrically applied their own standards for evaluating evidence of task demand, and in so doing have misrepresented the available evidence regarding demand on both sides of the debate.

When standards are applied equally to evaluating all research, the strength of the evidence against top-down effects is hardly as convincing as F&S suggest. For example, they rely on work by Durgin et al. (2011a) to argue that task demand plagues studies of top-down effects. However, this research fails to meet the standards F&S establish (see arguments 4.1.3, 4.4.3). Specifically, Durgin et al. (2011a) did not replicate the focal effect – that valuable objects appear closer than less-valuable objects. The researchers had participants toss a beanbag at a target. Study 1 found that task instructions that may manipulate task demand – to come closest to or to hit the target – produced different distributions of tossing behaviors. Study 2 found that financial incentives to come closest did not affect tosses to a neutral target. It is an erroneous inference to suggest these results undermine the conclusion that object value influences distance perception. Because they do not replicate the effect of object value, the conclusion that demand *explains* the effect of object value on distance perception collapses under the standards F&S put forth. Though informative, the findings are actually irrelevant to the study of object value on distance perception. The conclusion that demand effects *can* exert an influence on perceptual experience should not be confused for evidence that demand *does* exert an influence.

F&S's suggestions for how to mitigate task demand confuse rather than clarify the issue, because they introduce new possibilities for demand rather than serve as demand-free comparison conditions. They encourage researchers to mitigate demand by telling participants that external factors do not affect perception, which instead impacts compliance and honesty of reported perceptual experiences. Likewise, they encourage researchers to offer participants alternative explanations for experimental conditions that are not relevant to the predictions – for example, that props held during the task are meant to assist with balance rather than affect body size. F&S assert that alternative cover stories free participants of the influence of demand. We strongly contest this claim. Alternative cover stories do not remove the opportunity to guess hypotheses, nor do they eliminate the possibility that participants will amend their responses in accordance with their conjectured suppositions. They simply introduce new task demands. Researchers who attempt to overcome the pernicious effects of demand by manipulating expectations use a technique that is inherently flawed.

Instead, we offer five of our own published techniques as effective methods for overcoming demand. First, we use *accuracy incentives* to limit the likelihood of response bias (e.g., Balcetis & Dunning 2010; Balcetis et al. 2012). In a binocular rivalry paradigm where two different images were presented simultaneously, participants reported what they saw (Balcetis et al. 2012). We associated one image type (e.g., letters) with specific positive point values and another (e.g., numbers) with specific negative point values. We converted points into raffle tickets for monetary prizes. To ensure that participants truthfully reported their perceptual experience, we offered participants an accuracy incentive. If participants correctly reported *all* of the visual information they saw (e.g., both the letter and number that appeared), their score increased by some undefined amount; if they did not, their score decreased. Despite the incentive to respond accurately, participants primarily reported seeing only the images associated with reward and very infrequently reported experiencing both percepts. Moreover, if participants were strategically choosing to report percepts in ways that maximized payoff, we should have and could have found evidence for inhibition in addition to facilitation effects. However, participants were no less likely to report perceiving images associated with the loss of points relative to a baseline condition.

Second, we use *counterintuitive behavioral responses* as dependent measures, such that even if participants guessed the purpose of the study, they would not know how to respond to support the hypotheses (e.g., Balcetis & Dunning 2010; Stern et al. 2013). For example, participants estimated the distance to a desirable (chocolates) or undesirable object (dog feces) by moving themselves to stand a set, referent distance away from it (Balcetis & Dunning 2010, Study 3b). If the chocolate appeared closer, paradoxically, participants would need to stand farther away from it to match the set distance. That our predictions required they stand farther away from chocolates and closer to dog feces is a counterintuitive response that participants do not expect, reducing the likelihood of a demand effect.

Third, we conduct studies using *between-subjects designs* (e.g., Balcetis & Dunning 2010; Cole & Balcetis 2013). Participants cannot know that our hypotheses predict they will stand farther from chocolates if randomly assigned to the chocolate condition, and closer to feces if randomly assigned to that condition. Task demand is less likely to pertain when participants lack half of the information necessary to conjecture what the hypotheses expect.

Fourth, we use *double-blind hypothesis testing* in which both participants and experimenters are unaware of the assigned condition (e.g., Cole & Balcetis 2013). For example, when participants drank juice sweetened with either sugar or Splenda before estimating the distance to a location, participants, experimenters, and the graduate student training experimenters and analyzing data were unaware of drink-type. This reduced if not negated the impact of task demand, response bias, and experimenter bias.

Finally, we *dissociate participants' perceptual experiences* from manipulations that could be influenced by demand (e.g., Cole & Balcetis 2013). In a test of object construal on distance perception, participants tossed a beanbag with the intent to hit a picture frame that held a $100 bill or was empty. No financial reward or outcome was tied to the toss itself. Thus, the toss served solely as a behavioral measure of perceptual experience. By dissociating the perceptual measure from attainment of the reward, we reduced the likelihood that the measure reflected task demands.

These five techniques limit the opportunity for task demand and improve on F&S's suggestion to combat demand by manipulating expectations. We encourage researchers – including F&S – to apply the same rigorous standards not only to the analysis of studies that seek to provide confirmatory evidence for top-down effects, but also to studies that seek to provide disconfirmatory evidence.

The folly of boxology

doi:10.1017/S0140525X15002630, e231

Diane M. Beck and John Clevenger
Department of Psychology, University of Illinois, Urbana-Champaign, IL 61820.
dmbeck@illinois.edu
http://becklab.beckman.illinois.edu/
jcleven2@illinois.edu

Abstract: Although the authors do a valuable service by elucidating the pitfalls of inferring top-down effects, they overreach by claiming that vision is cognitively impenetrable. Their argument, and the entire question of cognitive penetrability, seems rooted in a discrete, stage-like model of the mind that is unsupported by neural data.

Firestone & Scholl (F&S) argue for the impenetrability of perception by cognition in part by relegating many well-established top-down effects to peripheral, trivial changes in sensory processing akin to willfully closing ones' eyes. Drawing on a large neuroscience literature, we argue that many of the class of effects the authors dismiss are neither trivial nor peripheral, but rather reflect complex tuning changes implemented across multiple levels of the visual hierarchy. More generally, the authors' arguments betray a conception of the mind that is difficult to square with modern neuroscience. In particular, they view the visual brain as containing a sensory/input module, a perception module, and a cognition module. Although carving the mind into these boxes has, in the past, provided a convenient construct for thinking about vision, this construct is unsupported by what we now know about the organization of the visual brain.

Although primary visual cortex (V1) is a well-established input to other cortical areas, it is only an input in the sense that the vast majority of visual input to the cortex first passes through it. Importantly, however, it is not encapsulated with respect to subsequent areas or task-demands. Activity in V1 has been shown to modulate in response to attention (e.g., Motter 1993; O'Connor et al. 2002), task demands (e.g., Li et al. 2006), interpretation, (Hsieh et al. 2010; Kok et al. 2012; Roelfsema & Spekreijse 2001; van Loon et al. 2015) and also to vary as a function of conscious experience (e.g., Lee et al. 2005; Wunderlich et al. 2005). Further, re-entrant activity in V1, long after the initial feedforward sweep of information, appears to be necessary for conscious perception (Pascual-Leone & Walsh 2001). These data indicate that V1 is not so much an early stage of vision but rather part of ongoing, dynamic network of feedforward and feedback activity. Even the lateral geniculate nucleus (LGN), a subcortical region that passes information from the retina to V1, tracks changes in attention (O'Connor et al. 2002) and conscious experience (Wunderlich et al. 2005). In other words, early visual regions are better thought of as part of a dynamic network of visual areas rather than an encapsulated sensory stage. Thus, even though it might be conceptually convenient to cleanly separate sensation from perception, the empirical data suggest that this dichotomy does not actually exist in the brain.

Even if, for the sake of argument, we equated early visual areas with an input or sensation module, other neural data do not support the authors' intuition that attention works by first modulating those early stages (i.e., their *peripheral* attention effects). The anatomical pathways through which attention initially modulates visual processing are relatively late in the visual hierarchy (Baldauf & Desimone 2014; Buffalo et al. 2005; Esterman & Yantis 2010; Moore & Armstrong 2003). In fact, the levels at which attention networks interface with visual cortex correspond more closely to a "perception box"; that is, later visual areas where activity seems to correlate more robustly with conscious perception and behavior (Hung et al. 2005; Leopold & Logothetis 1996; Logothetis & Schall 1989; Walther et al. 2009). Attention effects thus seem to feed back from later areas to earlier areas, not proceeding from input to perception as the authors' box model implies. Indeed, the authors' entire concept of *peripheral* attention effects is unsubstantiated by neural data.

Finally, the authors' intuition that attention effects are functionally similar to trivial changes in input also does not actually accord with the data. Although we agree that closing one's eyes is a trivial effect of cognition on vision, it is a mistake to take this intuition about the peripheral nervous system (i.e., the eye) and apply it to the brain. Instead of merely enhancing or gating the processing of stimuli (a simple gain model), attention also appears to cause large-scale tuning changes across the visual hierarchy. For example, when people view video clips and search for either humans or vehicles, the tuning of the entire ventral visual cortex, including V1, shifts toward the attended category and away from unattended categories (Çukur et al. 2013). Similarly, given identical stimuli, neurons in V1 can flexibly change whether they carry information about collinearity of lines or bisection of parallel lines depending on which task has been cued (Gilbert & Li 2013). Attention can even enhance orientations that are not actually present in the display (i.e., on either side of a target in orientation-tuning space) to more optimally discriminate the target from distractors (Scolari et al. 2012).

These attention effects are nothing like turning off the lights. Rather than simply gating or enhancing input, much of the neural hardware responsible for vision flexibly changes its function in complex ways depending on the goals of the observer. It is hard to imagine better evidence for the cognitive penetrability of vision than this.

The authors do seem aware of some of the neural data discussed here; they even admit that attention can work in "rich and nuanced ways," and "change the content of perception rather than merely influence what we focus on" (sect. 4.5, para. 6). Curiously, here they momentarily appear to back off their main thesis and restrict their criticisms to "*peripheral* sorts of attention – involving simple changes in which locations, features, or objects we focus on" (emphasis in original, sect. 4.5, para. 6). However, as we have discussed, the interactions among the cortical areas involved in vision are so extensive and result in such flexible representations throughout visual cortex that not only is it impossible to neatly separate sensation and perception, but also the concept of peripheral attention is rendered useless. Although these box models of the mind have great appeal and have facilitated both careful experimentation and fruitful theorizing in the past, the neural data are clear: It is time we move beyond a box model of the brain.

Tweaking the concepts of perception and cognition

doi:10.1017/S0140525X15002733, e232

Ned Block
Department of Philosophy, New York University, New York NY 10003.
ned.block@nyu.edu
http://www.nyu.edu/gsas/dept/philo/faculty/block/

Abstract: One approach to the issue of a joint in nature between perception and cognition is to investigate whether the concepts of perception and cognition can be tweaked to avoid direct, content-specific effects of cognition on perception.

Firestone & Scholl (F&S) give lip service to the claim that the real issue is that of a joint in nature between cognition and perception, but their focus in practice is on denying "true" top-down effects. For this reason they neglect the need to tweak the concepts of perception and cognition in order to home in on the joint—if there is a joint. I suggest a reorientation toward the question of whether direct, content-specific effects of cognition on perception can be eliminated by such tweaking.

I will explore this idea with regard to what is perhaps the most dramatic of the known top-down effects of cognition on perception. It has been shown that mental images can be superimposed on percepts, creating a combined quasi-perceptual state. Brockmole et al. (2002) used a "locate the missing dot" task in which the subject's task was to move a cursor to a missing dot in a 5-by-5 array. If a partial grid of 12 dots appeared briefly followed in the same place in less than 50 ms by another partial grid of 12 different dots (that stays on the screen until the response), subjects were able to fuse the two partial grids and move the cursor to the missing dot, remembering nearly 100% of the dots on the first array. (If the subject erroneously clicked on a space in which there was a dot in the first array, that was counted as forgetting a dot in the first array.) However, if the second array was delayed to 100 ms, subjects' ability to remember the first array fell precipitously (from nearly 12 dots down to 4 dots).

Brockmole et al. explored extended delays – up to 5 seconds before the second array appeared. The amazing result they found was that if the second array of 12 dots came more than 1.5 seconds after the first array had disappeared, subjects became very accurate on the remembered dots. (See Fig. 1, following.) Instructions encouraged them to create a mental image of the first array and superimpose it on the array that remained

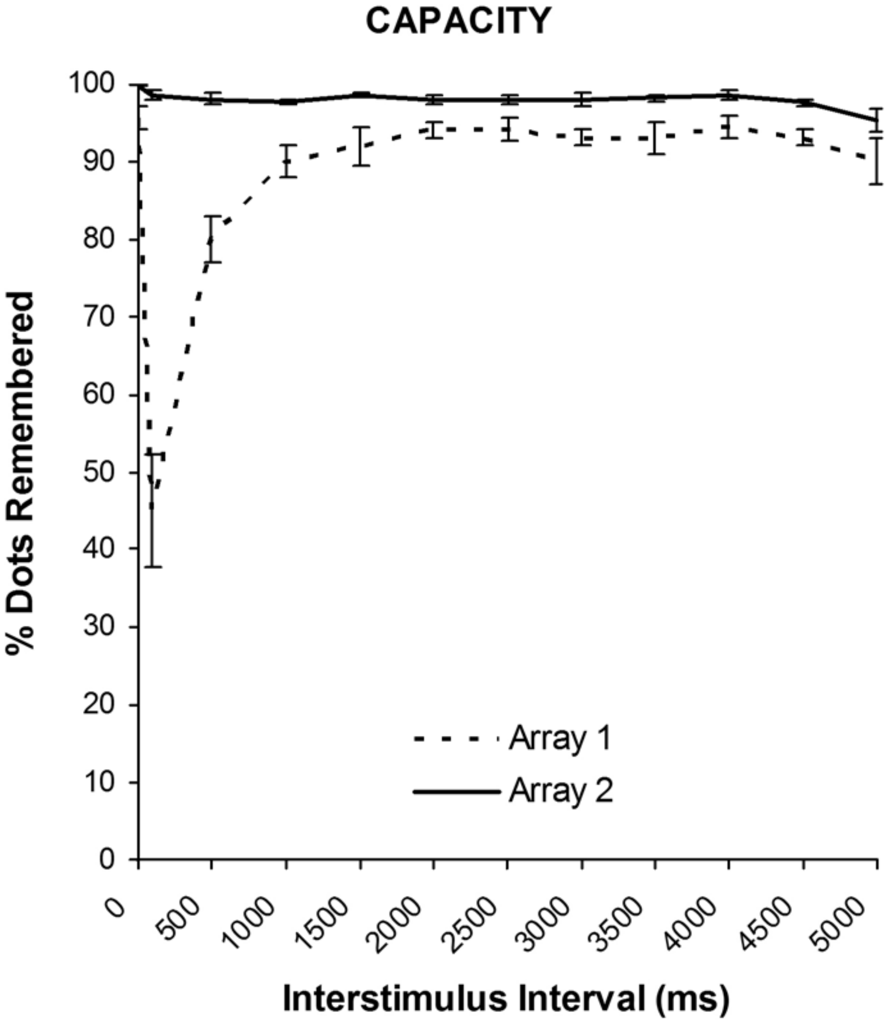


Figure 1 (Block). A 12-dot partial grid appears, then disappears. Another 12-dot partial grid appears – in the same place – up to 5 seconds later (i.e., after an interstimulus interval of up to 5 seconds). The second grid stays on the screen until the subject responds. The subject's task is to form a mental image of the first grid, superimposing it on the second grid so as to identify the dot missing from both grids. The dark line shows that subjects rarely click on an area that contains a dot on the screen. The dashed line shows that after a 1.5 second interval between the grids, subjects also rarely click on an area in which there was a dot in the remembered array. Thanks to James Brockmole for redrawing this figure.

on the screen. Brockmole et al. analyzed the subjects' errors, finding that subjects remembered more than 90% of the dots that they saw for up to 5 seconds. Further, the 1–1.5 seconds required to generate the mental image fits with independent estimates of the time required to generate a mental image (Kosslyn et al. 2006). Kosslyn and colleagues replicated this result with different methods (Lewis et al. 2011).

Those results demonstrate a direct content-specific effect of cognition on perception. And effects of this sort are involved in many of the phenomena used to criticize the idea of a joint in nature between cognition and perception. Many of the effects of language and expectation on perception that opponents of a joint in nature (e.g., Lupyan 2015a) cite seem very likely to involve superimposition of imagery on perception. For example, Lupyan notes that in a visual task of identifying the direction of moving dots, performance suffers when the subject hears verbs that suggest the opposite direction. Hearing a story about motion can cause motion aftereffects (Dils & Boroditsky 2010b), a point that enhances the plausibility that the result Lupyan notes involves some sort of combination of imagery and perception.

Another example: Delk and Fillenbaum (1965) showed that when subjects are presented with a heart shape and asked to adjust a background to match it, the background they choose is redder than if the shape is a circle or a square. As Macpherson (2012) points out, there is evidence that subjects are forming a mental image of a red heart and superimposing it on the cutout. Macpherson regards that as a case of "cognitive penetration." But if this is cognitive penetration, why should we care about cognitive penetration? Mental imagery can be accommodated to a joint between cognition and perception by excluding these quasi-perceptual states, or alternatively, given that imagery is so slow, by excluding slow and effortful quasi-perceptual states.

In expert perception – for example, wine-tasting, or a doctor's expertise in finding tumors in x-rays – we find another illustration of the benefit of using direct, content-specific, top-down effects of cognition on perception to tweak the category of perception. Opponents of a joint (Lupyan 2015a; Vetter & Newen 2014) often cite expert perception. But, as Fodor noted, the category of perception need not be taken to include perceptual learning (1983). The concept of perception can be shrunk so as to allow diachronic cognitive penetration in perceptual learning while excluding synchronic cognitive penetration.

Moving to a new subject: F&S's treatment of attention is inadequate. Oddly, they focus on spatial ("peripheral") attention and treat attention as akin to changing the input. However, Lupyan (2015a) is right that feature-based attention operates at all levels of the visual hierarchy (e.g., both orientations and faces) and so is not a matter of operation on input. Taking direction of motion as an example, here is how feature-based attention works: There is increased gain in neurons that prefer the attended direction; there is suppression in neurons that prefer other directions; and there is a shift in tuning curves toward the preferred direction (Carrasco 2011; Sneve et al. 2015). Although feature-based attention is directed by cognition, it may in many cases be mediated by priming by a sample of the feature to be attended (Theeuwes 2013). These effects cannot be eliminated by tweaking the concepts. But contrary to Lupyan (2015a), feature-based attention works by well-understood *perceptual mechanisms*, and there is not the slightest evidence that it directly modulates conceptual cognitive representations. This is a direct content specific effect of cognition on perception but what we know of the mechanisms by which it operates gives no comfort to opponents of a joint in nature between cognition and perception.

Finally, although I have been critical of F&S, I applaud their article for its wonderful exposé of the pitfalls in experimental approaches to the joint in nature between cognition and perception.

Acting is perceiving!

doi:10.1017/S0140525X15002460, e233

Rouwen Cañal-Bruland,[a] John van der Kamp,[b] and Rob Gray[c]

[a] *Institute of Sport Science, Department of Sport Psychology, Friedrich-Schiller-University Jena, 07749 Jena, Germany;* [b] *Department of Human Movement Sciences, Vrije Universiteit Amsterdam, 1081 BT Amsterdam, The Netherlands;* [c] *Human Systems Engineering, Arizona State University–Polytechnic, Mesa, AZ 85212.*

rouwen.canal.bruland@uni-jena.de **j.vander.kamp@vu.nl**
robgray@asu.edu
www.spowi.uni-jena.de/en/Sportpsychologie.html
https://vu-nl.academia.edu/johnvanderkamp
https://www.researchgate.net/profile/Rob_Gray/

Abstract: We challenge Firestone & Scholl's (F&S's) narrow conceptualization of what perception is and – most important – what it is for. Perception guides our (inter)actions with the environment, with attention ensuring that the actor is attuned to information relevant for action. We dispute F&S's misconceived (and counterfactual) view of perception as a module that functions independently from cognition, attention, and action.

Consider the following: An MLB pitcher throws a ball toward you that you need to hit. Presuming that most of our readers count as novices or, at best, less-skilled batters, then ample evidence indicates that – even if the pitcher performed the throw with the exact same movements, and hence would provide the exact same kinematic information as when he faced a professional pinch-hitter – you most likely would not (be able to) attune to and use the same information as the expert (Mann et al. 2007). The events with which you interact, the pitcher's throwing motion and the ball's flight, are identical and thus induce the very same patterns in the optic array, yet the information you attune to for guiding your batting action would be crucially different from the information the expert attunes to and uses. To paraphrase F&S, this is *the* perfect example for no change in visual input but a dramatic change in visual perception. It underlines our point that perception can be understood only by considering its primary purpose: to guide our actions in and interactions with the environment (Gibson 1979/1986; see also Cañal-Bruland & van der Kamp 2015).

Sticking with the batting example, there is no doubt that you and the professional pinch-hitter differ with respect to movement skills or action capabilities. Consequently, if the ball were to approach you and the expert player at the exact same speed, then the different movement skills would impose differential temporal demands on you with regard to initiating the batting response – that is, you would likely need to initiate the movement earlier than the professional. At the same time, the difference in movement skills, such as the temporal sequencing of the step and the swing, would require that – even though the exact same kinematic patterns were visually available to both – you and the expert player attune to and use different information (e.g., you may have to rely relatively strongly on early movement kinematics, whereas the expert may even be able to exploit early ball flight information).

Put in more general terms, individual action capabilities fundamentally constrain the way actors encounter their environment, and hence, the requirements differ with respect to the information actors need and attend to. In this respect, attention is an actor's attunement to the information that specifies the relevant properties of the environment–actor interaction. This attunement has multiple characteristics: (a) it is task-specific, (b) it is prone to differ within actors across multiple encounters, and (c) it is also prone to differ among actors (e.g., with different levels of skill). These characteristics follow because individuals guide their (inter)actions by accomplishing certain perceptions (Gibson 1979/1986). Let's elaborate on the characteristics of attunement to further specify our argument:

Attunement is task-specific. Referring back to the batting example: Although when batting, the professional pinch hitter is better attuned than you are to the kinematic information contained in the movements of the pitcher, this baseball player may not be better attuned than you are when he is watching and waiting on-deck before entering as a pinch hitter (Dicks et al. 2010; Milner & Goodale 2008; van der Kamp et al. 2008).

Attunement is prone to differ within actors across multiple encounters. The validity of this argument can be illustrated for different timescales. For example, on their long route to expertise, professional batters converge toward more-specific information available from a pitcher's throwing actions and a ball's flight. That is, with development, learning, or both, those actors who ultimately reach expert levels have learned to attend to information that is simultaneously uniquely specific to and adaptive for their superior batting actions. Within ecological psychology, this process of change is referred to as *education of attention* (Jacobs & Michaels 2007). Also, on a shorter timescale, such as across multiple throws by the same pitcher within one inning, convergence to more specific information may occur. For example, when a batter faces a new pitcher with a different pitching style, the professional batter may adapt to this style within a few pitches simply by shifting attention to a different invariant within the kinematic information contained in the pitcher's throw – even when the throwing kinematics remain stable over successive pitches.

Attunement is prone to differ among observers. As indicated above, in various domains such as sports and music, ample evidence demonstrates that given their action capabilities, experts differ as concerns the ability to pick up the most reliable information compared with their less-skilled counterparts (Huys et al. 2009; Mann et al. 2010). For example, expert baseball batters more frequently and effectively use information provided by the pitcher's arm angle and the ball's rotation (Gray 2002; 2010). Again, these differences in attention across individuals emerge as a result of differences in some combination of adaptation, learning, and development (and perhaps predisposition).

To reiterate, we argue that differences in perception always go together with differences in attention induced by the requirements for adaptive (inter)actions of actors within their environment. This inseparability of perception and action (reflected in attention), in our view, challenges F&S's core claims (1) that perception is an isolable entity or process that functions independently from cognition, attention, and action, and hence (2) that top-down processes do not impinge on perception. F&S's extremely narrow view of what perception is, namely bottom-up processing that results in percepts, neglects the fundamental question what perception is for. Yet, answering the question of what perception is without considering what perception is for naturally – as is the case in F&S's target article – leads to isolated percepts, which may trigger a few contrived issues but are rather marginal, if not meaningless, for understanding situated behaviours.

Attention alters predictive processing

doi:10.1017/S0140525X15002472, e234

Andy Clark
School of Philosophy, Psychology, and Language Sciences, University of Edinburgh, Edinburgh EH8 9AD, United Kingdom.
andy.clark@ed.ac.uk
http://edin.ac/1tqX3sO

Abstract: Firestone & Scholl (F&S) bracket many attentional effects as "peripheral," altering the inputs to a cognitive process without altering the processing itself. By way of contrast, I highlight an emerging class of neurocomputational models that imply profound, pervasive, nonperipheral influences of attention on perception. This transforms the landscape for empirical debates concerning possible top-down effects on perception.

A key move in Firestone & Scholl's (F&S's) impressive, skeptical review of putative top-down effects on perception is to bracket many attentional effects as "peripheral" – as working by altering the inputs to a cognitive process rather than by altering the processing itself. The deflationary thought (developed in sect. 4.5 of the target article) is that many attentional effects may be rather like turning one's head towards some interesting stimulus. Head-turning may indeed be driven top-down in ways sensitive to what the agent knows and intends. But in such cases, it seems that knowledge and intentions simply alter (via action) what the visual system gets as input. This idea is fully compatible with subsequent visual processing occurring in an inflexible, encapsulated (knowledge-impermeable) manner. To show genuine (interesting, nontrivial) top-down effects upon perception therefore requires showing that such effects impact not just the inputs but *the processing itself*, and do so in ways sensitive to what the agent (or better, I'll argue, the system) knows. F&S (sect. 5, para. 2–3) appeal to this possibility to "deflate" empirical evidence (such as Carrasco et al. 2004) suggesting that selective attention alters what we see even in cases where no overt action (such as head-turning or visual saccade) is involved. Selective attention, the authors concede, may indeed enhance specific aspects of the visual input. But such effects may be functionally similar to turning one's head, altering inputs that are then processed using a modular, inflexible system.

It is certainly conceivable that attention might always or mostly operate in just the blunt fashion that F&S suggest. But recent work on the so-called Bayesian brain, and (especially) related neurocomputational proposals involving predictive processing, already suggest a different, and much richer, mode of influence. That work (see Bastos et al. 2012, and reviews in Hohwy 2013; Clark 2013; 2016) depicts online perception as subject to two distinct but interacting forms of top-down influence. The first involves simple prediction: Downwards and lateral connections are said to be in the business of predicting the ongoing stream of sensory stimulation, using a generative model that aims to construct the inputs using stored knowledge. That constructive process is answerable to the evidence in the sensory stream, and mismatches yield prediction error signals that drive further processing. But that answerability is itself subject to *another* form of knowledge-based top-down influence. This second mode of influence involves the so-called precision-weighting of select sensory prediction error signals. Precision-weighting mechanisms alter the influence (postsynaptic gain) of specific prediction error signals so as to optimize the relative influence of top-down predictions against incoming sensory evidence, according to changing (mostly subpersonal) estimations of the context-varying reliability of the sensory evidence itself (see Feldman & Friston 2010).

This suggests a far richer model of attentional effects than the one suggested by F&S. Attention, thus implemented, is able to alter the balance between top-down prediction and bottom-up sensory evidence at every stage and level of processing. That same process enables such systems to reconfigure their own effective architectures on a moment-by-moment basis. They can do so because altering the precision-weighting on specific prediction error signals alters the influence that one neural area or processing regime has on other (specific) neural areas or processing routines (see, e.g., den Ouden et al. 2010). The result is a highly flexible cognitive architecture in which attention (construed as precision estimation) controls the flexibility.

In searching for possible top-down effects on perception, the authors lay great stress on the impact (or lack of impact) of top-level agentive reasons and intentions on the fine-grained course of processing. But this emphasis, from the Bayesian/Predictive Processing perspective, is also potentially misleading. In such models, the impact of top-down processing is best identified with the joint impact of priors and context-variable precision estimations. The question of how such priors and precision estimations interact with top-level agentive reasons and intentions is then a further (complex and important) one. But regardless of how that story goes, it seems likely that much of the knowledge that impacts online processing will be sub-agential, involving unconscious estimations of the reliability, in context, of various kinds of information for the task at hand.

Suppose, to take a concrete example, that I decide to look for my car keys on a crowded desktop. Suppose further that that decision automatically generates a cascade of altered estimations concerning the reliability or salience of different prediction error signals calculated as the processing unfolds, and that these alterations impact what we see as we scan the desktop. Once thus "seeded" by my top-level decision to seek out the keys, the unfolding flow of these effects could be automatic, reflecting nothing *further* about my intentions or goals. But wouldn't that still amount to a legitimate and interesting case of a top-down effect on perception? Such cases (memory-based variants of which have been experimentally explored by Nobre et al. 2008) may look superficially like ones in which attention "merely" enhances certain inputs. But if attention alters precision estimations at many levels of processing, that begins to look much more like the kind of profile that F&S demand – a profile in which the *processing itself* systematically changes in response to changing top-level goals and intentions (for some early explorations in these ballparks, see Feldman & Friston 2010; Vossel et al. 2014 – see also Gazzaley & Nobre 2012).

F&S (sect. 4.5) do allow that not all attentional effects need be (in their sense) "peripheral," and that Bayesian/Predictive Processing accounts may already suggest deeper modes of influence. But to take this caveat seriously should, I think, fundamentally alter the shape of the debate. It should alter the shape of the debate because attention (conceived as the process of optimizing precision estimations) then emerges as a deep, pervasive, and entirely *non-peripheral* player in the construction of human experience – one that acts not as a simple spotlight but as a subtle tool capable of altering the flow and detail of online processing in multiple ways.

The practical upshot is that F&S's downgrading of most attentional effects to simple alterations in what is given as input should be treated with caution. That downgrading is inconsistent with powerful emerging neurocomputational frameworks that depict attentional effects as reaching deep into the underlying processing regime. The conceptual contrast here is stark enough. But teasing these scenarios apart will require new experimental paradigms and some delicate model-comparisons.

The myth of pure perception

doi:10.1017/S0140525X15002551, e235

Gerald L. Clore and Dennis R. Proffitt
Department of Psychology, University of Virginia, Charlottesville, VA 22904-4400.
gclore@virginia.edu **drp@virginia.edu**
www.people.virginia.edu/~gc4q
www.people.virginia.edu/~drp

Abstract: Firestone & Scholl (F&S) assume that pure perception is unaffected by cognition. This assumption is untenable for definitional, anatomical, and empirical reasons. They discount research showing nonoptical influences on visual perception, pointing out possible methodological "pitfalls." Results generated in multiple labs are immune to these "pitfalls," suggesting that perceptions of physical layout do indeed reflect bioenergetic resources.

A commonly held notion in the nineteenth century was that experience was composed of immutable "pure sensations." This idea, which William James described as "as mythical an entity as the Jack of Spades" (James 1890, p. 236), was abandoned early in the twentieth century. Firestone & Scholl's (F&S's) arguments follow from a similar anachronistic assumption that there exists a state of immutable pure perception.

The notion of pure perception is as untenable today as pure sensation was 100 years ago. Retinal ganglia cells terminate in at least a dozen different brain areas, and feedforward and feedback reciprocal neural connections are ubiquitous in the visual system. According to our contemporary understanding from neuroscience, the brain cannot be carved at its joints so as to isolate pure perception anatomically and functionally.

Perception is, by definition, a *meaningful* awareness of one's environment and one's perspective on it. Lacking access to cognition, pure perception would be devoid of meaning and consist instead of an absolute associative agnosia in which the conceptual recognition of environmental objects and events is entirely absent. F&S argue that pure perception can be carved out of the flow of mental experience, but they are vague about what its constituents would be. It cannot include conceptual entities, by their account, because concepts fall on the other side of the perception–cognition divide. Being conceptual, even familiar objects cannot serve to scale space in pure perception.

F&S's definition of *pure perception*, however, is even more restrictive; they place the meaning of perceived distance outside the purview of pure perception. In fact, to perceive distance, the angular units of visual information must be transformed into linear units appropriate for measuring extent. This transformation requires geometry and a ruler. The meaning of a distance therefore is specified by the magnitude that it subtends on a ruler.

Following Gibson (1979), Proffitt and Linkenauger (2013; Proffitt 2013) proposed that spatial layout is scaled by perceptual rulers derived from the affordances associated with intended actions. As was certainly the case for Gibson, we view affordances as being neither top-down nor cognitive. F&S, however, lump action influences on perception in with cognitive ones, and they provide no alternative account for the meaningful scaling of spatial layout.

The studies F&S critiqued are an unrepresentative sample of the relevant literature. For example, the first experiments that found an influence of wearing a backpack on people's perception of hill slant were conducted 20 years ago (e.g., Proffitt et al. 1995). Since then, numerous studies from multiple labs have found evidence for bioenergetic scaling of perceived spatial layout using designs that avoid all of the methodological pitfalls emphasized in the target article. In those studies, researchers too approaches such as:

The "biological backpack." Taylor-Covill and Eves (2016) used an individual differences design, devoid of experimental manipulations, to show that percent body fat – what they called a biological backpack – was positively correlated with the perceived slant of stairs. Moreover, by assessing participants over time, they found that changes in body fat, but not fat-free mass, were associated with changes in apparent slant.

Bioenergetic scaling of walkable distances. Also employing an individual differences design, Zadra et al. (2016) found that physical fitness (measured as VO_2 max at blood lactate threshold) was inversely correlated with perceived distance.

Energy expenditure correlation. White et al. (2013) used a treadmill and a virtual environment viewed in a head-mounted display to decouple the relationships between energy expended while walking and the optically specified distance traversed. They found that increasing energy expenditure while keeping constant optically specified distance led to increased distance judgments.

Such findings show that extents and slants are scaled by the bioenergetic costs of walking and ascending. The theoretical generalization from the original backpack studies has thus been repeatedly upheld, but F&S claim that such results merely reflect experimenter demands. We find their evidence unconvincing, for the following reasons:

The studies purporting to control for demand were themselves demand-ridden (e.g., Durgin et al. 2012). Participants were warned not to respond the way others did, inflating hill slant estimates to please experimenters. Not surprisingly, those participants responded by giving low estimates of slant compared with the estimates of unwarned participants. Far from eliminating implicit demand, such instructions introduce explicit demand. The literature on "debiasing" (e.g., Schwarz 2015) indicates that such warnings tend to bias, rather than debias, judgments.

Participants in Durgin et al. (2009) were told that the backpack contained equipment to measure muscle potential. An elaborate story ensured that participants were alerted to the irrelevance of their experience of the backpack for judging slant. Additionally, a noisy cooling fan on the backpack served as a constant reminder of its irrelevance. Rather than eliminating demand, the elaborate effort to make the backpack both salient and distinct ensured that its heaviness would be segmented from other bioennergetic cues when making slant estimates. The results seem foreordained by the procedures. The findings likely reflect not only the odd backpack manipulation, but also the fact that the "hill" was only a 2-meter-long ramp. This hill did not afford the opportunity to walk more than a step or two, rendering any bioenergetic costs of wearing a backpack irrelevant (Proffitt 2009).

To assess awareness, experimenters (e.g., Durgin et al. 2009) first asked participants to consider the backpack heaviness and then asked for their hypotheses. The literature on assessing awareness has long warned against such procedures (e.g., Bargh & Chartrand, 2000; Dulany 1962), because participants cannot reliably distinguish hypotheses elicited by being asked questions from hypotheses actively generated during the experiment. To avoid such contamination, recommended methods involve carefully designed funnel interviews (e.g., Dulany 1962), which were not used in this research.

Other studies have found that replenishing depleted glucose can lower slant estimates (Schnall et al. 2010), but F&S suggest this occurs not because changes in resources affect perception, but because added glucose empowers participants to resist experimenter demands. We are not aware of any data supporting glucose effects on susceptibility to demand. More important, such suggestions cannot account for the results of multiple new studies, only some of which we cited earlier, which are simultaneously immune to the criticisms of F&S and supportive of the perceptual effects they attack.

ACKNOWLEDGMENTS

This work was partially supported by grant number BCS-1252079 from the National Science Foundation to Clore.

Bottoms up! How top-down pitfalls ensnare speech perception researchers, too

doi:10.1017/S0140525X15002745, e236

Anne Cutler[a] and Dennis Norris[b]

[a]*The MARCS Institute, Western Sydney University, Locked Bag 1797, Penrith South, NSW 2751, Australia;* [b]*MRC Cognition and Brain Sciences Unit, Cambridge CB2 2EF, United Kingdom.*

a.cutler@westernsydney.edu.au **dennis.norris@mrc-cbu.cam.ac.uk**

www.westernsydney.edu.au/marcs/our_team/researchers/professor_anne_cutler

www.mrc-cbu.cam.ac.uk/people/dennis.norris

Abstract: Not only can the pitfalls that Firestone & Scholl (F&S) identify be generalised across multiple studies within the field of visual perception, but also they have general application outside the field wherever perceptual and cognitive processing are compared. We call attention to the widespread susceptibility of research on the perception of speech to versions of the same pitfalls.

Firestone & Scholl (F&S) review an extensive body of research on visual perception. Claims of higher-level effects on lower-level processes, they show, have swept over this research field like a "tidal wave." Unsurprisingly, other areas of cognitive psychology have been similarly inundated.

Auditory perception and visual perception are alike in the questions they raise about the interplay of cognitive processing with sensory and perceptual analysis of a highly complex and externally determined input; and like visual perception, the processes underlying auditory perception have been well mapped in recent decades. Many of the features of the vision literature F&S note have direct auditory counterparts, such as highly intuitive demonstrations (the McGurk effect, whereby auditory input of [b] combined with visual articulation of [g] produces a percept of [d]; McGurk & MacDonald 1976), control of peripheral attention shift (the "cocktail-party phenomenon," attending to one interlocutor in a crowd of talking people), or novel pop-out effects, for example, for certain phonologically illegal sequences (Weber 2001).

Auditory and visual perception differ, however, not only in sensory modality but also in the input domain: Visual signals play out in space; auditory input arrives across time. The temporal input dimension has had implications for how the equivalent debate in the speech perception literature has played out; it has proved natural and compelling to treat the question of modularity as one concerning the temporal order of processing – has the bottom-up processing order (e.g., of speech sounds before the words they occur in) been compromised? Studies in which ambiguous speech sounds are categorised differently in varying lexically biased contexts (e.g., a [d/t] ambiguity reported as "d" before *-eep* but as "t" before *-eek* because *deep* and *teak* are words but *teep* and *deek* are not; Ganong 1980) were initially taken as evidence for top-down effects. (Note, however, that rather than tapping directly into perceptual processes, categorisation tasks may largely reflect metalinguistic judgments, as per F&S's Pitfall 2.) This work prompted follow-ups showing, for example, stronger lexical effects in slower responses (Fox 1984), and no build-up of effects with an ambiguous sound in syllable-final rather than syllable-initial position (McQueen 1991); these temporal arguments suggested a response bias account (F&S's Pitfall 3).

In general, F&S's Pitfall 1 (a confirmatory approach) has been the hallmark of much of the pro-top-down speech perception literature. Most of that literature takes the form of a catalogue of findings that are consistent with top-down effects but are not diagnostic. There is frequently little evidence that alternative feedforward explanations have been considered. One of the few exceptions comes from the study of compensation for coarticulation, a known low-level process in speech perception whereby cues to phonetic contrasts may be weighted differently depending on immediately preceding sounds. An influential study by Elman and McClelland (1988) reported that interpretation of a constant word-initial [t/k] ambiguity could be affected by the lexically determined interpretation of a constant immediately preceding word-final [s/sh] ambiguity (whether it served as the final sound of *Christmas* or *foolish*). Pitt and McQueen (1998) reasoned that if this compensation was a necessary consequence of the lexical effect (rather than of transitional probability, as they argued; that is, an artefact as per F&S's Pitfall 4), then if there is no compensation effect there should be no lexical effect either. With the [s/sh] occurring instead in words balanced for transitional probability (*juice*, *bush*), the word-initial [t/k] compensation disappeared; but the lexical [s/sh] effect remained. Such studies (testing disconfirmation predictions) are, however, as in the vision literature, vanishingly rare.

In our account in this journal of these and similar sets of studies (Norris et al. 2000), we concluded, as F&S do for the vision literature, that there was then no viable evidence for top-down penetrability of speech-perception processes. In a more recent review (Norris et al. 2016) we reach the same conclusion regarding current research programs in which similar claims have been reworded in terms of prediction processes ("predictive coding").

Priming effects (F&S's Pitfall 6) are more or less the bread and butter of spoken-word recognition research, so that psychological studies tend to preserve the memory/perception distinction; but in an essentially separate line of speech perception studies, from the branch of linguistics known as sociophonetics, the distinction has in our opinion been blurred. Typical results from this literature include listeners' matching of heard tokens to synthesised vowel comparison tokens being influenced by (a) telling participants the speaker was from Detroit versus Canada (Niedzielski 1999), (b) labeling participants' response sheets "Australian" versus "New Zealander" (Hay et al. 2006), or (c) having a stuffed toy kangaroo or koala versus a stuffed toy kiwi in the room (Hay & Drager 2010). In fairness to these authors, we note that they do not propound large claims concerning penetration of cognition into primary auditory processing (they interpret their results in terms of reference to listening experience). It seems to us, however, that a rich trove of possible new findings could appear if researchers would adopt F&S's advice and debrief participants, then correlate the match responses to debriefing outcomes.

A comprehensive and thorough review of a substantial body of research (with potentially important implications for theory) is always a great help to researchers – and especially useful if it uncovers new patterns such as, in this case, a systematic set of deficiencies. But in the present article, a service has been performed for researchers of the future as well, in the form of a checklist against which the evidence for theoretical claims can be evaluated. Only research reports that pass (or at least explicitly address) F&S's six criteria can henceforth become part of the serious theoretical conversation. As we have indicated, these criteria have application beyond visual perception; at least speech perception can use them too. Thus, we salute F&S for performing a signal service to the cognitive psychology community. Bottoms up!

Attention and multisensory modulation argue against total encapsulation

doi:10.1017/S0140525X1500254X, e237

Benjamin de Haas,[a,b] Dietrich Samuel Schwarzkopf,[a,b] and Geraint Rees[a,c]
[a]*Institute of Cognitive Neuroscience, University College London, WC1N 3AR London, United Kingdom;* [b]*Experimental Psychology, University College London, WC1H 0AP London, United Kingdom;* [c]*Wellcome Trust Centre for Neuroimaging, University College London, WC1N 3BG London, United Kingdom.*
benjamin.haas.09@ucl.ac.uk **s.schwarzkopf@ucl.ac.uk** **g.rees@ucl.ac.uk**
https://www.ucl.ac.uk/pals/research/experimental-psychology/person/benjamin-de-haas/
https://sampendu.wordpress.com/sam-schwarzkopf/
http://www.fil.ion.ucl.ac.uk/~grees/

Abstract: Firestone & Scholl (F&S) postulate that vision proceeds without *any* direct interference from cognition. We argue that this view is extreme and not in line with the available evidence. Specifically, we discuss two well-established counterexamples: Attention directly affects core aspects of visual processing, and multisensory modulations of vision originate on multiple levels, some of which are unlikely to fall "within perception."

Firestone & Scholl (F&S) argue there is no good evidence for cognitive penetration of perception, specifically vision. Instead, they propose, visual processing is informationally encapsulated. Importantly, their version of encapsulation goes beyond Fodor's original proposal "that at least *some* of the background information at the subject's disposal is inaccessible to at least some of his perceptual mechanisms" (Fodor 1983, p. 66). Their hypothesis is much more ambitious: "perception proceeds without any direct, unmediated influence from cognition" (sect. 5, para. 1). We will refer to this view as *total* encapsulation.

One possible counterexample to total encapsulation is *multisensory* modulation. For example, sounds in rapid succession can induce the illusory reappearance of visual flashes (Shams et al. 2000). Such reappearances increase objective sensitivity for visual features of the flash (Berger et al. 2003) and are linked to individual structure and function of primary visual cortex (de Haas et al. 2012; Watkins et al. 2006). Waving one's hand in front of the eyes can induce visual sensations and enable smooth pursuit eye movements, even in complete darkness (Dieter et al. 2014). The duration of sounds can bias the perceived duration of concurrent visual stimuli (Romei et al. 2011), and sensitivity for a brief flash increases parametrically with the duration of a co-occurring sound (de Haas et al. 2013a). The noise level of visual stimulus representations in retinotopic cortex is affected by the (in)congruency of co-occurring sounds (de Haas et al. 2013b). Category-specific sounds and visual imagery can be decoded from early visual cortex, even with eyes closed (Vetter et al. 2014), and the same is true for imagined hand actions (Pilgramm et al. 2016). At the same time, the location of visual stimuli can bias the perceived origin of sounds (Thomas 1941), and a visible face articulating a syllable can bias the perception of a concurrently presented (different) syllable (McGurk & MacDonald 1976). F&S argue that multisensory effects can be reconciled with total encapsulation. The inflexible nature and short latency of such effects would provide evidence they happen "within perception itself," rather than reflecting the effect of "more central cognitive processes *on* perception" (sect. 2.4, para. 1). However, multisensory effects have different temporal latencies and occur at multiple levels of processing, from direct cross-talk between primary sensory areas to top-down feedback from association cortex (de Haas & Rees 2010; Driver & Noesselt 2008). They may further be subject to attentional (Navarra et al. 2010), motivational (Bruns et al. 2014), and expectation-based (Gau & Noppeney 2015) modulations. Therefore, evidence regarding a strictly horizontal nature of multisensory effects seems ambiguous at best. If total encapsulation hinges on the hypothesis of strictly horizontal effects, this hypothesis needs to be clearer. Specifically, what type of neural or behavioural evidence could refute it?

A second, perhaps more definitive, counterexample is *attentional* modulation of vision. F&S acknowledge that attention can change what we see (cf. Anton-Erxleben et al. 2011; Carrasco et al. 2008) and that these effects can be under intentional control. For example, voluntary attention can induce changes in the perceived spatial frequency (Abrams et al. 2010), contrast (Liu et al. 2009), and position (Suzuki & Cavanagh 1997) of visual stimuli. Withdrawal of attention can induce perceptual blur (Montagna et al. 2009) and reduce visual sensitivity (Carmel et al. 2011) and sensory adaptation (Rees et al. 1997). Nevertheless, F&S argue for total encapsulation. On such an account, attention would not interfere with visual processing per se but with the *input* to this process, "similar to changing what we see by moving our eyes" or "turning the lights off" (sect. 4.5, para. 4).

Attention-related spatial distortions and changes in acuity have been linked to effects on the spatial tuning of visual neurons (Anton-Erxleben & Carrasco 2013; Baruch & Yeshurun 2014). Receptive fields can shift and grow towards, or shrink around, attended targets (e.g., Womelsdorf et al. 2008). Such effects go beyond mere amplitude modulation and can provide important evidence regarding their locus. In a recent study (de Haas et al. 2014), we investigated the effects of attentional load at fixation on neuronal spatial tuning in early visual cortices. Participants performed either a hard or an easy fixation task while retinotopic mapping stimuli traversed the surrounding visual field. Importantly, stimuli were identical in both conditions – only the task instructions differed. Performing the harder task, and consequently having to pay less attention to the task-irrelevant mapping stimuli (Lavie 2005), yielded a blurrier neural representation of the surround, as well as a centrifugal repulsion of population receptive fields in V1-3 (pRFs; Dumoulin & Wandell 2008). Importantly, this repulsion in V1-3 was accompanied by a centripetal *attraction* of pRFs in the intraparietal sulcus (IPS), perhaps because the larger receptive fields in IPS specifically encode the attended location (Klein et al. 2014). Critically, retinotopic shifts merely inherited from input modulations cannot trivially explain such *opposing* shifts, because any such effect should be the same (or very similar) across the visual processing hierarchy.

How can one reconcile these findings with total encapsulation? We can think of only one way: redefining *visual processing* in a way that excludes processing associated with retinotopic tuning of visual cortex but includes feedback processes from multisensory areas (as outlined in our second paragraph above). Such a re-defintion seems hard to reconcile with the widespread evidence that visually tuned neuronal populations in occipital cortex are involved with visual processing. We instead argue that attentional and multisensory modulations are inconsistent with *total* encapsulation, and at least here, the line between cognition and perception is blurred. F&S concede that accepting these exceptions to total encapsulation would be far less revolutionary than many of the claims they attack. We second their demand to back up extraordinary claims with rigorous evidence and applaud the standards they propose. Many effects they discuss may indeed fail to meet these standards. But precisely because attentional and multisensory effects are well established, total encapsulation itself strikes us as an extraordinary claim that is not supported by the available evidence.

ACKNOWLEDGMENTS
This work was supported by a research fellowship from the Deutsche Forschungsgemeinschaft (BdH; HA 7574/1-1), European Research Council Starting Grant 310829 (DSS, BdH), and the Wellcome Trust (GR).

How cognition affects perception: Brain activity modelling to unravel top-down dynamics

doi:10.1017/S0140525X15002757, e238

Martin Desseilles[a,b,c] and Christophe Phillips[a]
[a]*Cyclotron Research Centre, University of Liège B30, B-4000 Liège, Belgium;*
[b]*Clinique Psychiatrique des Frères Alexiens, B-4841 Henri-Chapelle, Belgium;*
[c]*Department of Psychology, University of Namur, B-5000 Namur, Belgium.*
martin.desseilles@unamur.be **c.phillips@ulg.ac.be**
http://mentalhealthsciences.com/index_en.html
http://www.cyclotron.ulg.ac.be/cms/c_15006/fr/christophe-phillips

Abstract: In this commentary on Firestone & Scholl's (F&S's) article, we argue that researchers should use brain-activity modelling to investigate top-down mechanisms. Using functional brain imaging and a specific cognitive paradigm, modelling the BOLD signal provided new insight into the dynamic causalities involved in the influence of cognitions on perceptions.

We were surprised to read that Chaz Firestone and Brian Scholl (F&S) consider cognition and perception as purely psychological mechanisms. Like the vast majority of professional neuroscientists worldwide, we consider that cognitions and perceptions are governed by specific patterns of electrical and chemical activity in the brain, and are thus essentially physiological phenomena. We therefore consider that cognitions and perceptions are embodied. Accordingly, we propose that brain activity best explains top-down mechanisms.

First, double-task *paradigms*, as proposed by Schwartz et al. (2005), allow testing focused attention at the centre of a screen (monitored by eye-tracking), in which attentional load at the centre while detecting various shapes and colours in a continuous stream of letters – L's and T's – can be varied simply by changing the instructions.

Simultaneously, irrelevant stimuli are presented at the periphery to activate specific areas of the ventral visual pathway (e.g., visual area 4 – V4) independently of the main attentional task.

Double-tasking enables the determination of whether certain top-down effects are related exclusively (or not) to intentional control – for example, during a directed attentional task, as F&S demonstrate. In our studies, we used divided attention in a double-task design that simultaneously varied the attentional load centrally and the perception of irrelevant coloured stimuli at the periphery (Desseilles et al. 2009; 2011).

Second, during our tasks we measured *blood-oxygen-level dependant (BOLD) signals* elicited by neuronal activity throughout the brain. Signal modulations resulting from task effects reflect changes in regional activity, and hence changes in brain perceptions: in other words, sensations.

This strategy differs considerably from the standard method of directly asking subjects about how their perceptions vary with performed tasks. Verbal self-reports involve several biases, such as social desirability bias. The result is a subjective, post hoc impression of what was meant to be an objective measure of perception. In contrast, brain imaging provides a practically real-time measure of brain perception that avoids the desirability bias as well as subjective self-reports and consequent interpretability problems, for both subjects and researchers. Accordingly, we consider perception a neuronal activity that can be measured objectively. The issue of whether the BOLD signal reflects neuronal activity alone does not affect this approach. In fact, we agree with the authors that between the input (subject stimulation) and the output (subject's self-report), if the internal representations are psychological only, it would be impossible to pinpoint the top-down influence of cognition on perception. We therefore propose that brain activity measures should be used to unravel the mechanisms of top-down dynamics (Desseilles et al. 2011). Uncontrollable and highly variable phenomena call for precise measures.

Third, brain activity *modelling* is performed in a series of steps. The first step is to identify brain areas that show significant activity modulation by the task (Friston et al. 2007). The second step is to map functional connectivity in terms of psychophysiological interactions (Gitelman et al. 2003). Although this approach does not allow determining causality between areas, it tests interactions between psychological factors (here, attentional load) and physiological factors (here, brain activity) (Desseilles et al. 2009). After these exploratory steps, causal relationships between identified sets of brain regions can be determined by generative mechanistic modelling of brain responses, also called dynamic causal modelling (Friston et al. 2003). In this framework, concurrent hypotheses on brain functioning are expressed by different connectivity models, and the optimal model for a population can be identified from the subjects' brain activity (Desseilles et al. 2011; Penny et al. 2010).

Fourth, the cognitive charge used in our tasks depended on the different *populations* that participated in the studies, including healthy controls and depressed patients (Desseilles et al. 2009; 2011). Importantly, the inclusion of pathophysiology sheds new light on top-down connectivity, and investigations of perception and cognition should consider mental illnesses that directly affect cognition. Similarly, a paradigmatic shift has taken place, from psychological studies of personality to clinical case studies, for instance, by Phineas Gage. In view of all of the above, we argue that brain activity modelling should be the method of choice for unravelling top-down dynamics.

Oh the irony: Perceptual stability is important for action

doi:10.1017/S0140525X15002666, e239

Frank H. Durgin
Swarthmore College, Swarthmore, PA 19081.
fdurgin1@swarthmore.edu
http://www.swarthmore.edu/SocSci/fdurgin1/publications.html

Abstract: I review experiments in which drinking a sugarless drink causes some participants who have low blood sugar from fasting to give lower slant estimates. Ironically, this only occurs to the extent that they believe that they have received sugar and that the sugar was meant to make the hill look shallower; those who received sugar showed no similar effect. These findings support the hypothesis that low blood sugar causes greater participant cooperation – which, in combination with other experimental details, can lead participants to make judgments that can either seem to support the effort hypothesis or contradict it. I also emphasize the importance of perceptual stability in the perception of spatial layout.

The target article contains an excellent prescriptive exposition for scientific sanity checks. But it also argues there is an a priori impenetrability that challenges us to consider anew the relationship between perceptual experience and cognitive attribution.

When a person is asked to put on a heavy backpack in an experiment, experimenters can probe their perceptual experience of space by multiple methods. For example, in the original report of their backpack effect Bhalla & Proffitt (1999), argued that verbal judgments reflected (distortions of) conscious perception, whereas haptic matching tasks represented (undistorted) unconscious perception. When we (Durgin et al. 2009) simply provided an explanation for the backpack ("it contains EMG [electromyography] equipment so we can monitor your ankle muscles"), the verbal effect went away. Moreover, later investigations of haptic matching tasks (Durgin et al. 2010; 2011b; Li & Durgin 2011; Shaffer et al. 2014) suggest that they are controlled by conscious perception (often the comparison of misperceived haptic orientation with misperceived hill orientation). In the end, the dissociations between the biasing of verbal and the non-biasing of haptic matching appeared to reflect a dissociation between biases in judgment (attributional effects) and perceptual stability.

Some have summarized our work on experimental demand as showing that the effects of effort (say) can be eliminated by beliefs, but our data go much further than that. We realized that in studies in which blood sugar was manipulated quite carefully (Schnall et al. 2010), the experimental manipulations purporting to show that lower blood sugar led to higher slant estimates confounded several uncontrolled contaminants. The blood sugar manipulation (supplying a drink that did or did not contain sugar to participants in a state of low blood sugar from fasting) was always followed by a cognitive task, such as a Stroop task simply to pass the time (might this juxtaposition matter?). When replicating this procedure, we (Durgin et al. 2012) found that this sequence led participants to look for a relationship between the drink and the cognitive task, rather than the drink and the hill. Moreover, nearly all subjects who were not confident about being able to tell sugar from sweetener assumed the drink we gave them contained sugar (might this belief matter?). Finally, the studies of blood sugar all employed a heavy backpack manipulation prior to the exposure to the hill. The backpack manipulation, in particular, seemed like a puzzling choice in an otherwise relatively clean experiment. What if low blood sugar just makes people more cooperative with the demand of the backpack?

Shaffer et al. (2013) removed the intermediate cognitive task and used an EMG deception in conjunction with a cover story about the sugar/no-sugar drink itself (that it improves the EMG signal). Therefore, we could remove the experimental demand of the heavy backpack, leaving clever participants to notice that the experiment involved a drink manipulation followed (after time sitting to absorb the fluid) by a hill judgment. (To test for any residual backpack effect, for half of the subjects the experimenter carried the backpack; there was no effect of wearing the backpack on estimates of the hill slant.) Although we had also provided a cover story for the sweet drink, a quarter of our participants reported in a written survey that they thought the drink was intended to affect their perception of the hill.

Consider that these suspicious/insightful participants were evenly divided between those who had been given sugar in their drink and those who had not, whereas all had arrived at the lab in a state of low blood sugar due to fasting. But consider also that nearly everyone assumed that the drink contained sugar (even though in this experiment we told them that it contained only electrolytes). Given the belief that the drink contained sugar, participants who cooperated with experimental demand should have given lower estimates of the hill slant than those who did not cooperate with experimental demand. But if the effect of low blood sugar (previously fasting participants not actually given sugar) is to increase cooperation, then we should see the ironic effect that among those who deduced that the experiment concerned sugar, those who had received no sugar in their drinks should actually give lower estimates than everyone else – and that is exactly what happened.

Therefore, in the case where effort theory made a clear and direct prediction (low blood sugar should increase slant estimates), ironically, demand prediction won out: Low blood sugar increased cooperation with experimental demand, which, in this experimental context, led to lower estimates of hill slant. Note that the magnitude of these ironic effects was just as large as the effects observed in backpack studies, but in the opposite direction.

In addition to the methodological motivations emphasized in the target article (control experiments to test alternative hypotheses are important), in our pursuit of these questions a fundamental theoretical commitment has developed: In the perception of large-scale space, at least, we do seem to see large perceptual distortions that are fairly stable across individual and contexts (e.g., Durgin 2014). The best explanation of these distortions appeals to information-theoretic coding advantages (comparable to Huffman coding) that tend to perceptually magnify certain angular variables that are quite useful for the control of locomotor action. The crucial theoretical point here is that systematic and stable perceptual bias (causing the exaggeration of apparent slant and the underestimation of apparent distance) can be advantageous for action control, but that momentary destabilization of space perception by desire, fatigue, and so forth would tend to undermine the whole point of perception as a guide for action.

I have no a priori commitment to the impenetrability of perception. I thought the hill experiments were cool when I first learned of them. But the more one looks into them, the less likely they seem to have been correctly interpreted. Cognitive systems, in general, must walk a tightrope between stability and flexibility: We need to be able to take in new information without losing the old. It may be that perceptual systems tend actually to be self-stabilizing but that attribution systems are free to reflect more of the totality of our knowledge (boy, that hill seems steep!).

Gaining knowledge mediates changes in perception (without differences in attention): A case for perceptual learning

doi:10.1017/S0140525X15002496, e240

Lauren L. Emberson
Peretsman-Scully Hall, Psychology Department, Princeton University, Princeton, NJ 08544.
lauren.emberson@princeton.edu
https://psych.princeton.edu/person/lauren-emberson

Abstract: Firestone & Scholl (F&S) assert that perceptual learning is not a top-down effect, because experience-mediated changes arise from familiarity with the features of the object through simple repetition and not knowledge about the environment. Emberson and Amso (2012) provide a clear example of perceptual learning that bypasses the authors' "pitfalls" and in which knowledge, not repeated experience, results in changes in perception.

In an effort to establish that perceptual learning is not a top-down effect on perception, Firestone & Scholl (F&S) state that the "would-be penetrator is just the low-level input itself" (sect. 2.5). In other words, the authors claim, perceptual learning is not mediated by knowledge about the environment but is internal to the perceptual system, arising from repeated experience with a stimulus. Regardless of whether you agree with the guidelines and definitions F&S propose, Emberson and Amso (2012) provide a clear example of perceptual learning that bypasses the authors "pitfalls" and in which knowledge, not simply repeated experience, changes perception.

Emberson and Amso (2012) examined participants' visual percept of a complex Target Scene before and after a perceptual learning task. Participants were explicitly asked to color the scene according to their visual percept and no other factors, thereby meeting the criterion set out by F&S to dissociate perception from judgment (sect. 4.2, Pitfall 2).

Unbeknownst to participants, the Target Scene was ambiguous such that the novel object in the middle could be viewed as two disconnected objects or a single object behind an occluder (Fig. 1). The participants were selected based on having an initial disconnected percept and, after exposure, were categorized as Non-Perceivers or Perceivers depending on whether they changed their percept to connect the novel object.

In a crucial control, Emberson and Amso established that simple repeated exposure to the Target Scene did not result in perceptual learning. Instead, changes in visual percepts arose when participants sequentially viewed additional scenes containing the novel object in different orientations and visual contexts (Fig. 1: Consistent Scenes). Because the novel object was always occluded, participants needed to integrate visual information across the scenes to create a globally unambiguous representation of the novel object. Even with exposure to the consistent scenes, only half of the participants changed their percept of the Target Scene (i.e., become Perceivers). Having established that neither simple repetition of the Target Scene nor the novel object biases perception, what differentiates Perceivers and Non-Perceivers?

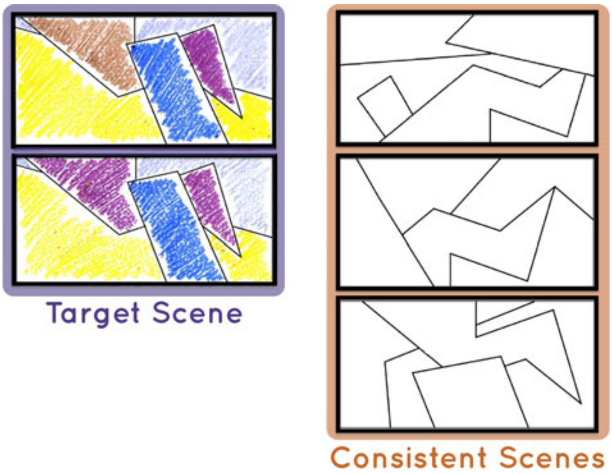


Figure 1 (Emberson). Complex scenes viewed during the perceptual learning task in Emberson and Amso (2012). Left panel: The Target Scene with two sample colorings depicting a typical disconnected percept (top) and a connected percept of the novel object (bottom). Participants all began with the disconnected percept (top). Participants viewed black-and-white scenes; color is for illustrative purposes only. Right panel: When participants viewed the Target Scene along with Consistent Scenes, half changed their percept to a connected percept (i.e., became Perceivers).

One possibility is that attention orienting (a measure of selective attention) distinguishes Perceivers from Non-Perceivers (sect. 4.5, Pitfall 5), but that is not the case. Emberson and Amso measured eye movements to the Target Scene.

Though there are subtle differences between Perceivers and Non-Perceivers, they are not significant at the group level and are dwarfed by the differences between participants in the control condition and those who viewed the Consistent Scenes. Therefore, attention orienting is most directly affected by whether information exists in the environment that can change perception and not a participant's visual percept per se.

It is neural activity in learning and memory systems that distinguishes Perceivers from Non-Perceivers. The pattern of activity observed for the hippocampus (Fig. 2) suggests that Perceivers are engaging learning and memory circuitry while viewing the Consistent Scenes to acquire knowledge about the novel object. Given that participants must integrate perceptual information across consistent scenes to obtain an unambiguous representation of the novel object, there is a natural match between this perceptual learning task and the hippocampus' neurobiological and computational abilities to encode conjunctive information (Frank et al. 2003). Indeed, recent work has demonstrated hippocampal involvement in associating information across sequentially presented episodes (Tubridy & Davachi 2011) and tracking the spatial relationship between an object and its context (Howard et al. 2011). Because that is the case for Perceivers only, the knowledge appears to bias conscious perception of the Target Scene, but through indirect means, because the circuitry is not active during viewing of the Target Scene.

What about Pitfall 6, Memory and Recognition? This pitfall is the least well supported by F&S. F&S's major argument is that top-down effects must be on "what we see [and not] how we recognize various stimuli" (sect. 4.6, para. 1), but this distinction is not motivated by, nor does it directly follow from, their definition of perception. Indeed, their argument is entirely definitional: "Given that visual recognition involves both perception and memory as essential but separable parts, it is incumbent on reports of top-down effects on recognition to carefully distinguish between perception and memory" (sect. 4.6.3). F&S do not establish, however, why this assumption must be made in the first place. Therefore, the door is left open for memory effects that avoid the remaining pitfalls to be clear evidence of cognitive penetration of the perceptual system according to their framework.

F&S also attempt to circumvent neuroimaging evidence for top-down effects by claiming that "nearly all brain regions subserve multiple functions" (sect. 2.2, para. 1). Regardless of whether you agree with F&S' line of reasoning, that criticism does not hold up for the current neural evidence. The argument made by Emberson and Amso (2012) is not that one can find activity in perceptual systems as evidence for a top-down effect but rather that changes in visual percepts arise from activity in learning and memory systems. For the F&S criticism of neural evidence to hold water, F&S would have to show that activity in these learning and memory systems has been associated with both perceptual and nonperceptual tasks (i.e., that the hippocampus subserves multiple functions). Evidence exists that regions of the medial temporal lobe are indeed crucial for perceptual as well as memory tasks (e.g., the perirhinal cortex; Graham et al. 2010). However, this work stops short of finding perceptual effects for the hippocampus proper, the region that differentiates Perceivers and Non-Perceivers. The hippocampus is one of the few regions of the brain that has not been found to subserve multiple

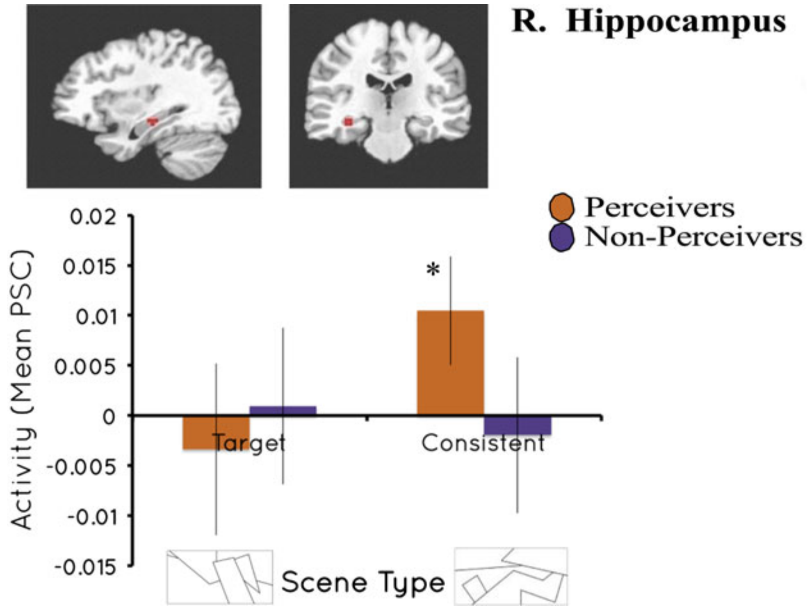


Figure 2 (Emberson). The pattern of activity in the right hippocampus responds to Consistent Scenes in Perceivers but not in Non-Perceivers. This region does not respond in either group to the presentation of the Target Scene. PSC, percent signal change; R. Hippocampus, right hippocampus; *, $p<0.05$.

functions, despite intense study for decades, and therefore, inferring that activity in this area is nonperceptual and related to the creation of memory or knowledge about the environment is well substantiated.

Finally, the remaining pitfalls are clearly circumvented in Emberson and Amso (2012): It does not have low-level stimulus differences to account for (Pitfall 4), is not subject to the El Greco fallacy (Pitfall 1), and has no difference in task demands (Pitfall 3).

Crossmodal processing and sensory substitution: Is "seeing" with sound and touch a form of perception or cognition?

doi:10.1017/S0140525X1500268X, e241

Tayfun Esenkaya[a,b] and Michael J. Proulx[b]

[a]*Department of Computer Science, University of Bath, Bath BA2 7AY, United Kingdom;* [b]*Department of Psychology, University of Bath, Bath BA2 7AY, United Kingdom.*

t.esenkaya@bath.ac.uk **m.j.proulx@bath.ac.uk**
http://www.bath.ac.uk/psychology/staff/michael-proulx/

Abstract: The brain has evolved in this multisensory context to perceive the world in an integrated fashion. Although there are good reasons to be skeptical of the influence of cognition on perception, here we argue that the study of sensory substitution devices might reveal that perception and cognition are not necessarily distinct, but rather continuous aspects of our information processing capacities.

We live in a multisensory world that appeals to all of our sensory modalities (Stein et al. 1988). Crossmodal perception is the norm, rather than the exception; the study of the senses as independent modules rather than integrated systems might give rise to misunderstandings of how the mind allows us to sense, perceive, and cognitively process the world around us (Ghazanfar & Schroeder 2006). Although some connections between the senses appear to be direct and low-level (Meredith & Stein 1986), others certainly are the product of learning and experience (Proulx et al. 2014). The authors of the target article provide an excellent framework and review for examining the (lack of) influence of cognition on perception, yet we think it would have been even stronger had they not dismissed crossmodal perception and cognition.

One area of crossmodal research that certainly connects perception and cognition is the study of sensory substitution and augmentation (an area of research that is cited in the fantastic textbook *Sensation and Perception* [Yantis 2013], unlike the top-down studies Firestone & Scholl [F&S] criticize). Sensory substitution devices allow someone with a damaged sensory modality (such as the visual system for the visually impaired) to receive the missing information by transforming it into a format that another intact sensory modality (such as the auditory [Meijer 1992] or somatosensory system [Bach-y-Rita & Kercel 2003]) can process. Learning plays a clear role in using these devices to access the otherwise inaccessible information. But is the ability to "see" with such a system better classified as perception or cognition? Although some have argued that it must be classified as seeing in the perceptual sense for psychology to be a science (Morgan 1977), we accept that this might be open to some debate.

Considering the distinction between perception and judgment in the target article, amongst other key issues, it seems that the authors might indeed debate whether sensory substitution allows for perception. Furthermore, there is evidence that long-term users of such devices who once had vision (who had acquired blindness) have visual imagery that is evoked immediately by the sounds they hear with a device, and therefore express that they have the perception of sight (Ward & Meijer 2010). Such cases might not be classified as an immediate crossmodal effect, such as those noted in the target article, so what might account for this ability instead? Any kind of information the sensory organs receive is meaningless because there is no inherent meaning in sensory information without experience and knowledge (Proulx 2011), and F&S acknowledge this point to some extent by allowing for unconscious inference to play a role in perception (von Helmholtz 2005). But if unconscious inference plays a role in perception, as it certainly does even in the perception of a red object, then there must be some role of recognition and judgment even for low-level perception.

The sharing of neural resources for the processing of perceived and imagined information (Klein et al. 2004; Kosslyn et al. 1999) is also suggestive of an interplay between perception and cognition, if not perhaps the idea that there is continuity between such steps of information processing, rather than categorical differences between them.

F&S provide an excellent checklist to carefully assess perception versus cognition that is useful regardless of the continual or categorical nature of these phenomena. For example, their discussion of a possible distinction between perception and judgment in the context of the El Greco fallacy provides a useful approach to assess how users are able to learn sensory substitution. The relative reliance on perception, judgment, response bias, memory, and recognition could be assayed at different points in training to provide a full profile of how the information is being transferred from one sense to another, and thus reveal the mechanisms of seeing with sound or touch. Novel methods such as sensory substitution, and related areas of crossmodal cognition including synesthesia, might provide crucial ways of examining perception and cognition in a new light (or sound).

Behavior is multiply determined, and perception has multiple components: The case of moral perception

doi:10.1017/S0140525X15002800, e242

Ana P. Gantman and Jay J. Van Bavel
Psychology Department, New York University, New York, NY 10003.
agantman@nyu.edu jay.vanbavel@nyu.edu
http://www.psych.nyu.edu/vanbavel/lab/index.html

Abstract: We introduce two propositions for understanding top-down effects on perception. First, perception is not a unitary construct but is composed of multiple components. Second, behavior is multiply determined by cognitive processes. We call for a process-oriented research approach to perception and use our own research on moral perception as a "case study of case studies" to examine these issues.

It is tempting to agree that top-down effects on perception (such as our own, *moral pop-out effect*; Gantman & Van Bavel 2014) constitute a radical reinterpretation of a foundational issue (Firestone & Scholl 2015b). Unfortunately, we cannot get excited about a notion of perception that excludes attention, inference, prediction, or expectations; that whittles the fascinating and broad domain of perception to sawdust. It is one thing for Firestone & Scholl (F&S) to argue that the entire field of visual neuroscience is irrelevant to understanding human perception. It is quite another to dismiss evidence of re-entrant processing precisely *because* it is well-established ("whatever one thinks of the relevance," sect. 2.2, para. 2). That makes F&S's cognitive penetrability argument circular.

F&S define perception to "encompass both (typically unconscious) visual processing and the (conscious) perceptions that result" (sect. 1, para. 3) as they "have the broader aim of evaluating evidence for top-down effects on *what we see* as a whole" (sect. 1.2, para. 2, emphasis theirs). Yet, they only consider perception that is separable from attention and occurs prior to – and independently from – memory, judgment, and social and physical context. It is difficult to understand how this definition might include "*what we see* as a whole." Moreover, their dismissal of unconscious inferences in vision, crossmodal effects, evidence of re-entrant processing in neuroscience, and changes in perceptual sensitivity over time, carves the mind at false joints. Their model of the mind seems to reify the administrative structure of psychology departments, manufacturing natural kinds of perception, cognition, and social processing.

The architecture of the mind does not recognize these distinctions. After perceptual input reaches the retina, multiple brain regions operate on the input – selecting the significant from the mundane (Lim et al. 2009), often by emotion (Anderson & Phelps 2001) or motivation (Egner & Hirsch 2005), and some via top-down re-entrant processes (Gilbert & Li 2013) – to construct perceptual experience. We suggest that the more pertinent question for future research is not *whether* top-down influences penetrate perception, but rather *which components* of the perceptual processing stream are sensitive to top-down influence. But perhaps this is a matter of preference.

Thankfully, the crux of F&S's argument lies in their empirical re-explanations of a handful of case studies. These are falsifiable. We invite readers to take a closer look at their case studies, which make strong claims about psychological process. We evaluate F&S's empirical claims regarding the moral pop-out effect and ask whether this exposes a fundamental problem with their case study approach.

We previously reported that moral words were more frequently detected than nonmoral words (matched for length and frequency), only when presented at the threshold for perceptual awareness (termed the *moral pop-out effect*, Gantman & Van Bavel 2014; 2015). The moral pop-out effect occurred over and above measured differences in valence, arousal, or extremity. We concluded that moral words were detected more frequently than control words.

F&S recently claimed that semantic memory must be solely responsible for the moral pop-out effect because the moral words were more related to each other than the control words were. As proof, they claimed to find "entirely analogous" fashion and transportation pop-out effects. Unfortunately there are fundamental flaws in the design and reporting of these studies that make it difficult for us to see how they could draw such strong conclusions.

First, there are some surprising problems with the experimental design of F&S's studies. For example, participants were never randomly assigned to experiments. Without this linchpin of experimental design, it is difficult, if not impossible, to draw any firm inferences about the similarities or differences between moral versus fashion/transportation pop-out effects.

Even more surprising, F&S's inferences rely on merely looking at the studies side by side and judging whether summaries of the results appear similar. Yet, there is good reason to believe these similar behavioral effects (moral vs. fashion/transportation word detection) arise by different processes even at the semantic level. We suspect that moral words are not explicitly encoded in semantic memory as moral terms or as having significant overlapping content. For example, *kill* and *just* both concern morality, but one is a noun referring to a violent act and the other is an adjective referring to an abstract property. Category priming is more likely when the terms are explicitly identifiable as being in the same category or at least as having multiple overlapping semantic features (e.g., pilot, airport).

Unfortunately, Firestone and Scholl (2015b) do not present the necessary data to evaluate whether a similar process (specifically, repetition priming, p. 42–43) underlies moral versus fashion/transportation experiments. Although they reported the relevant means for the fashion/transportation experiments – which support their claims that fashion/transportation words show repetition priming effects – they do not report the analogous means for their morality experiment That unfortunate omission seems particularly problematic because the moral perception case study – like all of the presented case studies – is presumably an argument about process.

Any observed behavior can be explained by multiple processes intervening between perceptual input and motor response. A single process rarely explains any behavior, and possible explanations are not always mutually exclusive. Accordingly, it is trivially true that semantic memory plays some role in the moral pop-out effect (how else would our participants know words like *kill* and *die*), yet this does not rule out that the motivational relevance of morality (or any motivationally relevant construct) boosts related content to awareness. Accordingly, we do not think morality is modular, either (Van Bavel et al. 2015). For example, we would expect a "pop-out" effect for any motivationally salient construct, such as food-related words when participants are food-deprived (Radel & Clément-Guillotin 2012). We suspect that other case studies suffer from this same failure to consider the multiple determinants of behavior.

Arguments about process – and especially mediation – must examine it directly. Otherwise, dismissing any effect with a 1:1 model of cause-to-behavior seems presumptuous (and even naive). These problems plague F&S's reinterpretation of the moral pop-out effect, and they may well be embedded in their reinterpretation of other case studies.

Action valence and affective perception

doi:10.1017/S0140525X15002605, e243

Walter Gerbino and Carlo Fantoni

Department of Life Sciences, Psychology Unit "Gaetano Kanizsa," University of Trieste, 34128 Trieste, Italy.

gerbino@units.it **cfantoni@units.it**
http://bit.ly/units_dsv_wg
http://bit.ly/units_dsv_cf

Abstract: Respecting all constraints proposed by Firestone & Scholl (F&S), we have shown that perceived facial expressions of emotion depend on the congruency between bodily action (comfort/discomfort) and target emotion (happiness/anger) valence. Our studies challenge any bold claim against penetrability of perception and suggest that perceptual theory can benefit from demonstrations of how – under controlled circumstances – observer's states can mold expressive qualities.

Referring to effects Firestone & Scholl (F&S) review in section 1.1, they claim: "There is in fact *no* evidence for such top-down effects of cognition on visual perception" (sect. 1.2, para. 1). F&S's thesis relies on a narrow concept of perception that – contrary to its standard meaning, inclusive of high-level perception – excludes memory-based attributes studied in recognition experiments (sect. 4.6). The narrowness of such a concept is implicit in their short list of "true" perceptual attributes ("colors, shapes, and sizes," sect. 1, para. 3) and related ostensive definitions (red, but not the price of the apple in the supermarket; the lightness difference in the target article's Fig. 1a). What about functional and expressive properties? Aren't they truly perceptual?

A risk of tautology. With respect to Pitfall 6 (sect. 4.6), F&S expose their thesis to the risk of being an unfalsifiable tautology. One could always preserve it from falsification by claiming that if any top-down factor influences *x* (any attribute provisionally taken as truly perceptual) then *x* does not belong to true "front-end" perception (Lyons 2011). This logical difficulty is overcome if perception is defined as an *explanandum* independent of its determinants (*explanantia*). By calling "*perception* (and, less formally, *seeing*)…both (typically unconscious) visual processing and the (conscious) percepts that result" (sect. 1, para. 3) and by taking front-end stimulus processing as the only truly perceptual stage, F&S mix the denial of the *explanandum* (any claimed top-down influence on perceptual experience is greatly exaggerated) and the rejection of not truly perceptual *explanantia*.

Categorical perception. The existence of categorical perception (Goldstone & Hendrickson 2010) seems to contradict F&S. This subfield is operationally defined by the warping of perceived similarity space, such that – once a categorical boundary is established – discrimination is better across categories than within categories. Categorical perception can be acquired during individual experience (Beale & Keil 1995), and it constitutes a paradigmatic case of observer-dependent, familiarity-mediated sensitivity change.

Tertiary qualities. Equally problematic for the thesis F&S defend is the domain of tertiary qualities as defined in the Gestalt literature (Köhler 1938; Metzger 1941; Sinico 2015; Toccafondi 2009). Consider the influence of observer's states on perceived facial expressions (not a revolutionary effect, given the naive idea that beauty is in the eye of the beholder). Despite being commonly conceived as subjective, expressive qualities are phenomenally objective (i.e., perceived as belonging to the object; Köhler 1929; 1938) and show a remarkable – though not exclusive – dependence on configural stimulus properties. Therefore, assessing observer-dependent effects on tertiary (in particular, expressive) qualities contributes significantly to perceptual science, which can tolerate, we believe, a circumscribed leakage of cognition into perceptual apartments, consistent with grounded cognition (Barsalou 2010; Kiefer & Barsalou 2013).

Penetrability of perceived facial expressions. To remain "empirically anchored" (sect. 4, para. 1), consider our studies on the effects of comfort and discomfort of bodily actions on perceived facial expressions of emotion. Using a novel motor action mood induction procedure (MAMIP), Fantoni and Gerbino (2014; Gerbino et al. 2014) demonstrated a congruency effect in participants who performed a facial emotion identification task along a happy-to-angry morph continuum after a sequence of visually guided reaches: A face perceived as neutral in a baseline condition appeared slightly happy after comfortable actions and slightly angry after uncomfortable actions. In agreement with F&S (sect. 4.2.3), we considered such evidence insufficient to claim that bodily action influences affective perception in a within-cycle fashion (better than "top-down," we deem): An action-induced transient mood might shift the point of subjective neutrality in identification by influencing only postperceptual (not perceptual) processing. In a subsequent study (Fantoni et al. 2016), we corroborated the perceptual nature of bodily action effects using an emotion detection task. Rather than response bias changes attributable to cognitive set shifts, MAMIP produced systematic, mood-congruent sensitivity changes in the detection of both positive and negative target emotions, with constant 0.354 d' increments ($p = 0.000$) in congruent (comfortable-happiness, uncomfortable-anger) over incongruent (uncomfortable-happiness, comfortable-anger) conditions at increasing percent emotion in the morph.

Referring to the final checklist F&S propose (sect. 5.2), our facilitation-by-congruency effect on emotion detection hold the following properties:

1. *It is robust* (relative to the Uniquely Disconfirmatory Predictions criterion) *and immune to the El Greco fallacy*; the detection threshold for happiness decreased about 2.2% after a sequence of congruent-comfortable than incongruent-uncomfortable reaches ($p = 0.001$), and the detection threshold for anger (in a different group of observers, to control for carryover effects) decreased about 7.4% after a sequence of congruent-uncomfortable than incongruent-comfortable reaches ($p = 0.003$);
2. It is operationalized by shifts of *performance-based measures* (d', absolute threshold, just noticeable difference) consistent over different paradigms (identification and detection);
3. It occurs even in the *absence of response bias shifts*, given that the response criterion *c* did not change significantly across congruency conditions for both happiness and anger detection;
4. It is dependent on the *observer's internal states* induced by motor actions unrelated to targets to be detected;
5. It is *independent of peripheral attention*; the performance improvement produced by uncomfortable actions, inducing high arousal, is selective (i.e., they improve anger, not happiness, detection);
6. It is *independent of memory* in the sense that the detection of subtle signs of happiness or anger in unfamiliar faces likely involves a general perceptual ability, independent of episodic memories of specific individuals.

To summarize, our evidence challenges the bold claim that cognition does not affect perception, supports a circumscribed penetrability of affective perception, and suggests *indirect affective priming* as a candidate mechanism by which observer's states mold perceptual properties experienced as phenomenally objective and yet loaded with meaning. If brains resonate with the similarity of motor and affective states (Niedenthal et al. 2010), bodily actions may behave as affectively polarized primes that preactivate the representations of emotionally related facial features, thus facilitating the *encoding* of features belonging to the same, rather than to a different, valence domain.

Carving nature at its joints or cutting its effective loops? On the dangers of trying to disentangle intertwined mental processes

doi:10.1017/S0140525X1500271X, e244

Robert L. Goldstone, Joshua R. de Leeuw, and David H. Landy

Department of Psychological and Brain Sciences, Program in Cognitive Science Indiana University, Bloomington, IN 47405.

rgoldsto@Indiana.edu **dhlandy@gmail.com**
josh.deleeuw@gmail.com
http://www.indiana.edu/~pcl/

Abstract: Attention is often inextricably intertwined with perception, and it is deployed not only to spatial regions, but also to sensory dimensions, learned dimensions, and learned complex configurations. Firestone & Scholl's (F&S)'s tactic of isolating visual perceptual processes from attention and action has the negative consequence of neglecting interactions that are critically important for allowing people to perceive their world in efficient and useful ways.

Firestone & Scholl (F&S) wish to draw a sharp distinction between attention acting to change the inputs to perception and top-down influences on perceptual processes proper. In the extreme example of a person coming to believe that important information is coming at them from the left and consequently moving their head to the left, this distinction is compelling. The belief causes an explicit head movement that thereby changes the input to perceptual processes. The perceptual processes may operate in a standard, unvarying way without their internal workings being directly affected by the belief. The problem with generalizing this kind of account is that attention influences perceptual processes at many stages and sometimes with a high degree of specificity. Because of the wide range of attentional effects – from centrally oriented modulations of percept selection, to adjustments of gain to certain features, to the extreme case of head-turning – the boundary between peripheral attention shifts and changes to perception is entirely unclear and potentially artificial. In our view, a better strategy for understanding the role of perception in human behavior is to consider the workings of larger attention-perception-learning-goal-action loops.

As F&S acknowledge, in some cases the action of attention is far more nuanced than moving one's head or opening one's eyes. Drawing a sharp distinction between cases in which attention is operating peripherally on the inputs to perception versus centrally on perception itself is counterproductive, leading to an unproductive debate about whether a particular process is part of "genuine" perception. The intellectual effort devoted to this debate would be more efficaciously applied to determining how different mental circuits coordinate to account for sophisticated and flexible behavior.

Attention is deployed not only to spatial regions, but also to sensory dimensions, learned dimensions, and learned complex configurations of visual components. Relegating attentional effects to occurring peripheral to perception is awkward given that these effects occur at many stages of perceptual processing. Neurophysiologically, when a particular cortical area for vision projects to a higher level, there are typically recurrent connections from that higher level back to the cortical area. These recurrent connections are particularly important when processing degraded inputs, and they serve to strengthen weak bottom-up signals (Wyatte et al. 2012). Treating feature-driven attention as distinct in kind from object or location-driven attention is equally problematic, because the apparent mechanisms underlying these processes are quite similar, and the impacts are entirely analogous.

Consider the evidence that attention can select complex learned configurations of relevance for a task. For example, the efficiency of search for conjunctions of visual parts gradually increases over the course of hours of training (Shiffrin & Lightfoot 1997). The learning of these complex visual features ought to be considered perceptual, rather than generically associative, based both on its neural locus in IT brain regions specialized for visual shape representation (Logothetis et al. 1995) and on behavioral evidence indicating perceptual constraints on the acquisition of these features (Goldstone 2000). For simpler auditory (Recanzone et al. 1993) and visual (Jehee et al. 2012) discriminations, even earlier primary sensory cortical loci have been implicated in perceptual learning. For both simple and complex perceptual learning, attention and perception are inextricably intertwined. Attention can be effectively deployed to subtle, simple discriminations and complex configurations only because perceptual processes have been adapted to fit task demands. Segregating these mechanisms into "attention" and "perception" and demanding that they be analyzed only separately ignores the key role of interacting systems in coordinating experience.

The interplay between perceptual processing and attention is even more striking in situations where perceptual learning leads to people separating dimensions that they originally treated as psychologically fused. For example, saturation and brightness are aspects of color that most people have difficulty separating. It is hard for people to classify objects on the basis of saturation while ignoring brightness differences (Burns & Shepp 1988). However, with training, dimensions that were originally fused can become more perceptually separated (Goldstone & Steyvers 2001), and once this occurs, it becomes possible for attention to select the separate dimensions. Perceptual changes affect what can be attended, and attention affects what is perceptually learned (Shiu & Pashler 1992). Given the intertwined nature of perception–attention relations such as these, it is appropriate to consider human behavior on perceptual tasks to be a product of an integrated perception–attention system.

Some might argue that perceptual learning occurs but is not driven by top-down influences such as expectations and goals. People are not yet generally able to directly implement the neural changes that they would like to have, but neurosurgery is only one way in which we can purposefully, in a goal-directed fashion, influence our perceptual systems. Athletes, musicians, coaches, doctors, and gourmets are all familiar with engaging in training methods for improving their own perceptual performances (Goldstone et al. 2015). For example, different music students will give themselves very different training depending on whether they want to master discriminations between absolute pitches (e.g., A vs. A#) or relative intervals (e.g., minor vs. major thirds). The neuroscientist Susan Barry provides a compelling case of the strategic hacking of one's own perceptual system: By presenting to herself colored beads at varying distances and forcing her eyes to jointly fixate on them, Barry caused her visual system to acquire the binocular stereoscopic depth-perception ability that it originally lacked (Sacks 2006).

To maintain that the intentions of learners only *indirectly* change perceptual processes only reinforces a distinction that conceals the consequential interactions between intention and perception that adapt perception to specific tasks. By analogy, when a person blows out a candle flame, does he or she blow it out directly or through an indirect chain of air pressure differentials, displacement of the flame away from the wick, and resulting temperature drop to the wick? Perseverating on this dubious distinction, or F&S's distinction between intentions directly versus indirectly affecting perception, risks neglecting the perception–

attention, perception–learning, and intention–perception loops that are critically important for allowing people to perceive their world in efficient and task-specific ways.

The anatomical and physiological properties of the visual cortex argue against cognitive penetration

doi:10.1017/S0140525X15002629, e245

Moshe Gur
Department of Biomedical Engineering, Technion, Haifa 32000, Israel.
mogi@bm.technion.ac.il

Abstract: We are consciously aware of visual objects together with the minute details that characterize each object. Those details are perceived instantaneously and in parallel. V1 is the only visual area with spatial resolution and topographical exactitude matching perceptual abilities. For cognition to penetrate perception, it needs to affect V1 image representation. That is unlikely because of the detailed parallel V1 organization and the nature of top-down connections, which can influence only large parts of the visual field.

Firestone & Scholl's (F&S's) target article is an important and timely critique of the present zeitgeist that supports fundamental, even detailed, influence of cognition on perception. The notion that a high-level mental event such as a mere semantic label can change basic percepts goes beyond having a revolutionary potential in changing our understanding of the modus operandi of the brain; it falls into the category of an "extraordinary claim." And extraordinary claims require extraordinary evidence. The notion of cognitive penetration into perception is extraordinary because, as I show below, no feasible physical route exists for such a penetration. Only by concentrating on psychophysical results – while totally refraining from suggesting *how* cognitive penetration can be instantiated – can the proponents of that view escape the realization of how extraordinary their claim is. Homeopathy, to take an extreme example, is recognized as making extraordinary claims because no plausible mechanisms for its effects are suggested.

My comments aim to complement the target article in pointing out that the anatomical and functional properties of the visual cortex preclude any significant, spatially specific cognitive penetration. To support this view I present two fundamental mechanisms related to the organization and functionality of the visual cortex: representation of space by primary visual cortex (V1) as evident by conscious perception, and the organization of top-down inputs.

Conscious perception of space is manifested first and foremost in the detailed, high-resolution, topographically precise depiction of spatial elements. When we see a face, for example, we perceive it as an integrated whole, but at the same time we are consciously aware of all of the minute spatial elements that make up that face. Indeed, it is our perception of details that enables us to discriminate between faces – in spite of their inherent similarity. Elsewhere, I show (Gur 2015) that because V1 is the only visual area that represents space at a resolution and topographical exactitude that is compatible with our perceptual abilities, its preintegration response patterns ("V1 map") must be the neural substance of image representation. From these preintegration response patterns, information converges in successive hierarchical stages, from V1 orientation-selective cells through downstream cortical areas V2, V4, and the inferotemporal cortex (IT). This successive convergence results in cells having increasingly large receptive fields (RFs) that cannot encode spatial details but may be selective to global features such as size, orientation, or category. At the pinnacle of the hierarchy we find anterior IT "object selective" and "face-selective" cells with very large (10°–50°) RFs responding to a considerable part of the visual field and requiring large (>5°) stimuli for their activation (Ito et al. 1995). Clearly, such cells cannot represent the fine spatial details that we are so sensitive to. Thus, in the visual cortex, objects in their full details are represented by V1 response patterns while the global information for each object that is required for comparing the acute image to its stored prototype ("recognition") is extracted by information-integrating cells at the various visual areas. Note that it is only the acute image, small or large, tilted or not, bright or dark, that is consciously perceived. All of the processes, from V1 preintegration activity patterns onward, that extract spatial global information are not perceived. We also note that the perceptual ability to transform individual spatial elements into holistic objects is instantaneous and parallel – which argues for a feed-forward, encapsulated perception.

Top-down inputs to V1. V1 is the only cortical area where space is represented at a high resolution, so for cognition to influence perception it must affect V1 space representation. However, top-down inputs to V1 do not match its topographic accuracy. Cognitive inputs, which originate, presumably, in nonvisual areas, go through multisynaptic, multiarea paths before reaching V1. For example, inputs originating at the prefrontal cortex reach first the IT cortex; some outputs from the IT cortex reach V1 directly and some reach it via the V4→V2→V1 route (Gilbert & Wu 2013). Almost all top-down information reaches V1 via layer 1 (see summary Fig. 1 in Gur & Snodderly 2008), where the spatial extent of its synaptic connectivity is rather spread out (Rockland 1994). Such connectivity rules out any direct, spatially circumscribed influence over V1 upper layers' cells representing the objects' details (Gur 2015; Gur et al. 2005). Even more important, the route into the visual cortex goes through the large RF cells of the IT cortex. Feedback from such cells, which are not sensitive to small spatial details, makes it impossible for top-down input to influence a selective part of the visual scene – such as a stroke on a Chinese character or an image of a dog surrounded by shoes (Fig. 2 of the target article). We can thus conclude that because cognition must reach V1 through IT large RF cells to end up in diffuse synaptic contacts in V1 layer 1, it can modulate processes, such as attention, affecting a large part of the visual field, but cannot target a small part of the visual scene while leaving other parts unchanged. Thus, the nature and organization of this fuzzy top-down information rule out any direct, spatially accurate influence.

On the neural implausibility of the modular mind: Evidence for distributed construction dissolves boundaries between perception, cognition, and emotion

doi:10.1017/S0140525X15002770, e246

Leor M. Hackel,[a] Grace M. Larson,[b] Jeffrey D. Bowen,[c] Gaven A. Ehrlich,[d] Thomas C. Mann,[e] Brianna Middlewood,[f] Ian D. Roberts,[g] Julie Eyink,[h] Janell C. Fetterolf,[i] Fausto Gonzalez,[j] Carlos O. Garrido,[f] Jinhyung Kim,[k] Thomas C. O'Brien,[l] Ellen E. O'Malley,[m] Batja Mesquita,[n] and Lisa Feldman Barrett[o,p]

[a]*Department of Psychology, New York University, New York, NY 10003;* [b]*Department of Psychology, Northwestern University, Evanston, IL 60208;* [c]*Department of Psychological and Brain Sciences, University of California, Santa Barbara, Santa Barbara, CA 93106;* [d]*Department of Psychology, Syracuse University, Syracuse, NY 13244;* [e]*Department of Psychology, Cornell University, Ithaca, NY 14853;* [f]*Department of Psychology, Pennsylvania State University, University Park, PA 16801;* [g]*Department of Psychology, The Ohio State University, Columbus, OH 43210;* [h]*Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN 47405;* [i]*Department of Psychology, Rutgers University, Piscataway, NJ 08854;* [j]*Department of Psychology, University of California, Berkeley, Berkeley, CA 94720;* [k]*Department of Psychology, Texas A&M University, College Station, TX*

77840; [l]Department of Psychology, University of Massachusetts, Amherst, Amherst, MA 01003; [m]Department of Psychology, State University of New York, Albany, Albany, NY 12222; [n]Center for Social and Cultural Psychology, University of Leuven, B-3000 Leuven, Belgium; [o]Department of Psychology, Northeastern University, Boston, MA 02115; [p]Department of Psychiatry and the Martinos Center for Biomedical Imaging, Massachusetts General Hospital/Harvard Medical School, Boston, MA 02114.
leor.hackel@nyu.edu gracelarson2017@u.northwestern.edu
bowen@psych.ucsb.edu gaehrlic@syr.edu tcm79@cornell.edu
blm266@psu.edu roberts.1134@osu.edu jeyink@indiana.edu
j.fetterolf@gmail.com fjgonzal@berkeley.edu carlosgarrido@psu.edu
jhkim82@tamu.edu tcobrien@psych.umass.edu
ellen.e.omalley@gmail.com mesquita@psy.kuleuven.be
l.barrett@neu.edu

Abstract: Firestone & Scholl (F&S) rely on three problematic assumptions about the mind (modularity, reflexiveness, and context-insensitivity) to argue cognition does not fundamentally influence perception. We highlight evidence indicating that perception, cognition, and emotion are constructed through overlapping, distributed brain networks characterized by top-down activity and context-sensitivity. This evidence undermines F&S's ability to generalize from case studies to the nature of perception.

Firestone & Scholl (F&S) rely on an outdated view of the mind to argue against top-down influences on perception. We highlight three of their assumptions that are untenable given contemporary neuroscience evidence: that the brain is modular, reflexively stimulus-driven, and context-independent. This evidence undermines their leap from critiques of individual studies to the conclusion that cognition does not affect perception.

The brain is not modular. F&S assume that the words *cognition* and *perception* refer to distinct types of mental processes ("natural kind" categories; Barrett 2009) localized to spatially distinct sets of neurons in the brain, sometimes called *modules* or *mental organs* (Fodor 1983; Gall 1835; Pinker 1997). As an intuition pump, the authors ask readers to "imagine looking at an apple in a supermarket and appreciating its redness" (sect. 1, para. 3). "*That* is perception," they suggest, compared with appreciating an apple's price, which they argue is "cognition." This modular view assumes that the brain's visual processing is "encapsulated" from nonperceptual influences. For example, F&S propose that context effects on perception are fully encapsulated within the visual system and therefore are not a meaningful example of top-down effects.

This approach promotes using phenomenology to guide scientific insight, which epitomizes *naive realism* – the belief that one's experiences reveal the objective realities of the world (Hart et al. 2015; Ross & Ward 1996). The distinctive experiences of seeing and thinking do not reveal a natural boundary in brain structure or function. The idea that the brain contains separate "mental organs" stems from an ancient view of neuroanatomy (see Finger 2001 for a history of neuroanatomy). Modern neuroanatomy reveals that the brain is better understood as one large, interconnected network of neurons, bathed in a chemical system, that can be parsed as a set of broadly distributed, dynamically changing, interacting systems (Marder 2012; Sporns 2011; van den Heuvel & Sporns 2013). These systems are domain general: Their interactions constitute mental phenomena that we consider distinct, such as perception, cognition, emotion, and action (for discussions, see Anderson 2014; Barrett 2009; Barrett & Satpute 2013; Lindquist & Barrett 2012; Pessoa 2014; Yeo et al. 2015). For example, Figure 1 displays a meta-analytic summary of more than 5,600 neuroimaging studies from the Neurosynth database (www.neurosynth.org; Yarkoni et al. 2011) showing brain "hot spots" that evidence a consistent increase in activity across a wide variety of tasks spanning the domains of perception, cognition, emotion, and action (for other evidence, see Yeo et al. 2015). Seemingly distinct mental phenomena are implemented as dynamic brain states, not as individual, static mental organs, violating the assumption that the mind has intuitive "joints."

The brain is not reflexively "stimulus driven." F&S assume that perception is reflexively driven by sensory inputs from the world that are commonly referred to as "bottom-up input" For example, they describe cross-modal effects and context effects on perception as occurring "reflexively" based on visual input alone. But again, neuroanatomy is inconsistent with claims of reflexiveness. Cortical cytoarchitecture is linked to information flow within the brain (see Barbas, 2015; for discussion, see Chanes & Barrett, 2016) and shows how representations of the

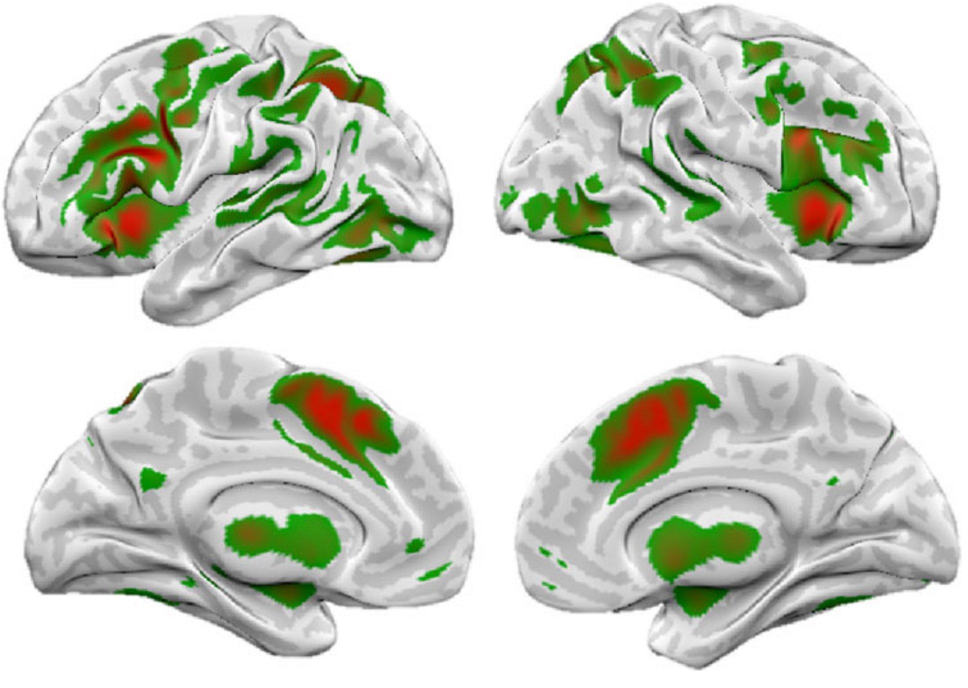


Figure 1 (Hackel et al.). Results of a forward inference analysis, revealing hot spots in the brain that are active across 5,633 studies from the Neurosynth database. Activations are thresholded at FWE $P<0.05$. Figure taken from Clark-Polner et al. (2016).

past, created in the vast repertoire of connectivity patterns within the cortex (referred to as "top-down" and colloquially called "memory" or "cognition"), are always involved in perception, and often dominate. Vision is largely a top-down affair (e.g., Gilbert & Li 2013). For example, by most estimates, only 10% of the synapses from incoming neurons to primary visual cortex originate in the thalamus, which brings sensory input from the retina; the remaining 90% of these synapses originate in the cortex itself (Peters 2002). Indeed, a bottom-up, reactive brain would be metabolically expensive and anatomically infeasible (e.g., see Sterling & Laughlin 2015). F&S dismiss top-down connections as irrelevant to their argument, because knowledge of anatomical connections is "common ground for all parties" (sect. 2.4, para. 2) and so these connections cannot be "revolutionary." The issue of their novelty is irrelevant, however: F&S are arguing a position that violates the functional architecture of the brain.

Top-down anatomical connections are consistent with a predictive brain that models the world through *active inference* (e.g., Bar 2007; Barrett & Simmons 2015; Clark 2013; den Ouden et al. 2012; Friston 2010; Rao & Ballard 1999). This process not only allows for, but also is *predicated on*, the existence of top-down effects. Specifically, the brain generatively synthesizes past experiences to continually construct predictions about the world, estimating their Bayesian prior probabilities relative to incoming sensory input. The brain then refines predictions accordingly. This means that top-down influences typically *drive* perception, and are constrained or corrected by incoming sensory inputs, rather than the other way around. Indeed, when humans and nonhuman animals change their expectations, sensory neurons change their firing patterns (e.g., Alink et al. 2010; Egner et al. 2010; Makino & Komiyama 2015; for a discussion, see Chanes & Barrett, 2016).

Although F&S acknowledge unconscious, *reflexive* inference in the visual system, they dispute the idea that "cognitive inferences" shape perception. Their distinction between reflexive visual inference and cognitive inference again advocates for a boundary that is rooted in naive realism, between reflex and volition (Descartes 1649/1989). It has long been known that the main distinction between automatic and controlled processing (or System 1 and System 2) is primarily phenomenological (for a discussion, see Barrett et al. 2004). Because the brain's control networks are involved in processing prediction error (applying attention to neurons to shape which inputs from the world are considered information and which are noise (Barrett & Simmons 2015; Behrens et al. 2007; Gottlieb 2012; Pezzulo 2012), they are always engaged to some degree; relative differences in activity should not be confused with "activation" and "deactivation" or "on" and "off"). It is more consistent with neuroanatomy to assume a continuum of brain modes, with one end characterized by brain states constructed primarily with prediction (e.g., phenomena called "memory," "daydreaming," "mind wandering," etc.), and the other end characterized by brain states where prediction error dominates (e.g., novelty-processing, learning, etc.), with a range of gradations in between. Evidence that the brain is active, not merely reactive, undermines the idea that perception is a bottom-up reflex isolated from cognition.

The brain is context-dependent. In discussing their "El Greco" fallacy, F&S implicitly assume that top-down effects would uniformly influence all elements in a visual field. This represents a fundamental misunderstanding of how the brain constructs Bayesian inferences, and in particular, reveals an underappreciated role of context. The authors argue that if an aperture looks more narrow when passing through it holding a wide rod (Stefanucci & Geuss 2009), then a second aperture that a participant adjusts to match the perceived width of the first (but is not required to pass through) should look similarly narrow, at least while one is holding the same rod. Following the authors' logic, these two distortions should cancel out, leaving no measurable impact of holding the rod on width estimates. However, the first aperture is meant to be passed through, whereas the second is not, and it is well known that requirements for action strongly shape moment-to-moment processing (Cisek & Klaska 2010). To assume that the top-down influence on width estimates would be the same and therefore cancel out under these distinct conditions suggests a misunderstanding of top-down effects.

Context shapes Bayesian "priors" that inform the brain's sensory predictions (Clark 2013; Friston 2010). As a consequence, sensory neurons behave differently when involved in contextually distinct perceptual tasks (Gilbert & Sigman 2007). Requirements for action (dictated by the task context) exert a top-down influence on the processing of visual information. In light of these considerations, research questions should shift from, "Are there top-down effects on perception?" in a global, undifferentiated way to "In what contexts and at what level in a hierarchy of predictions do different top-down effects emerge in a nuanced way?"

We agree with F&S that some studies of top-down effects may capture processes that are not traditionally categorized as perception (e.g., the impact of demand characteristics on judgment), and that studies designed for disconfirmation are an important part of theory testing. But the main thrust of their critique is based on folk categories of perception and cognition as reified in a modular, reactive, and context-insensitive brain. These assumptions, and the conclusion they support, are untenable given the considerable neuroscientific evidence that processing in the human brain is distributed, active, and exquisitely sensitive to context.

ACKNOWLEDGMENTS
This manuscript was supported by a US National Institute on Aging grant (R01AG030311), a US National Institute of Child Health and Human Development grant (R21 HD076164), and contracts from the US Army Research Institute for the Behavioral and Social Sciences (contracts W5J9CQ12C0049 and W5J9CQ11C0046) to Barrett. The views, opinions, and findings contained in this article are those of the authors and should not be construed as an official position, policy, or decision of the US National Institutes of Health or Department of the Army unless so designated by other documents.

Fundamental differences between perception and cognition aside from cognitive penetrability

doi:10.1017/S0140525X15002526, e247

Graeme S. Halford[a,b] and Trevor J. Hine[b]
[a]*School of Psychology, The University of Queensland, Brisbane 4067, Australia;* [b]*Behavioural Basis of Health, Menzies Health Institute Queensland and School of Applied Psychology, Griffith University, Brisbane 4122, Australia.*
g.halford@griffith.edu.au **t.hine@griffith.edu.au**
https://www.psy.uq.edu.au/directory/index.html?id=15#show_Research

Abstract: Fundamental differences between perception and cognition argue that the distinction can be maintained independently of cognitive penetrability. The core processes of cognition can be integrated under the theory of relational knowledge. The distinguishing properties include symbols and an operating system, structure-consistent mapping between representations, construction of representations in working memory that enable generation of inferences, and different developmental time courses.

There are fundamental and well-established differences between cognition and perception independent of cognitive penetrability. We will briefly outline properties that distinguish cognition from perception following Halford et al.'s (2010) contention that relational knowledge is the foundation of higher cognition.

Relational knowledge includes bindings between a symbol and an ordered set of elements. Thus, the relation "elephant larger_than mouse" entails a binding between the relation symbol "larger_than" and slots for a larger and a smaller element (in this example, filled with elephant and mouse, respectively). One of the core properties of higher cognition and not perception is the assignment to slots in a relational structure. Relational knowledge enables structure-based mapping between representations, including analogical reasoning, which is now established as fundamental to reasoning and the acquisition of many concepts (Gentner 2010). Halford et al. (2014) have shown that higher cognition is distinguished by structure-based mapping between representations, enabling inferences.

Symbols, which are fundamental to higher cognition, including both language and reasoning, are core properties of higher cognition. However symbols also require an operating system, which in turn depends on representation of structure. Therefore, to understand "cat eats mouse" we have to assign "cat" to the agent role and 'mouse" to the patient role. Thus, symbols depend on assignment to slots in a structure, a property that is shared with relational knowledge.

Inferences generated from representations that capture information in premises (Goodwin & Johnson-Laird 2005) are another core property of higher cognition: For example b larger_than c, a larger_than b can be integrated into a mental model in which a, b, and c are ordered (equivalent to monotonically_larger (a, b, c), permitting the reasoner to make the inference a larger_than c) . Inferences can be generated that are not perceived. Much of reasoning depends on mental models (Khemlani & Johnson-Laird 2012),

Involvement of working memory, which plays a distinctive role in forming representations, including assignment to slots in a coordinate system (Oberauer 2009). The transition from subsymbolic to symbolic cognition depends on assignment of elements to slots in a coordinate system in working memory (Halford et al. 2013).

Capacity limits of cognition are different from those of perception, and as humans, we are restricted to representing links between four variables (a quaternary relationship) or fewer in a single cognitive representation (Halford et al. 2005). No such limitation applies to perception.

Other properties of higher cognition have also been identified with relational knowledge (Halford et al. 2014). These properties include *compositionality*, which means preserving components in a compound representation: For example, in "John loves Mary," "John" retains its identity in the compound representation. Another fundamental property of higher cognition is *systematicity*, which means that representations are linked by common structure, so the ability to understand the proposition "John loves Mary" is linked to the ability to understand structurally similar propositions such as "Mary loves John" (Fodor & Pylyshyn 1988; Phillips & Wilson 2014).

Developmental time courses are another point of difference. The development of higher-order cognitive functions exhibits a much slower time course than the development of visual perception. The basic functions of human vision: acuity, contrast sensitivity, stereopsis, and the perception of two-dimensional and three-dimensional shape and size all have reached near-adult levels within the first year (Boothe et al. 1985). More important, even the most methodologically rigorous studies have shown that perceptual organisation following Gestalt principles (e.g., good continuation and good form) is clearly evident in the visual behaviour of 9-month-olds (Quinn & Bhatt 2005; Spelke et al. 1993). Finally, early work claimed that the most "cognitive" of percepts – Kanizsa's illusory contours – is evident in 7-months-olds (Berthenhal et al. 1980). Though later work has challenged this very young age (Nayar et al. 2015), what is not in dispute is that by 7 years of age, the functionality of human visual perception is indistinguishable from that of an adult.

A confused mass of literature has developed as a result of a lack of unifying concepts, and the consequent lack of conceptual coherence, or a paradigm. As a result, we lack tools for interpreting empirical findings. Motivation to reduce everything to a single set of core processes, common to a number of fields including cognitive development, blurs distinctions and means we fail to recognise core properties. The result is that we cannot slice nature at its joints. We can, however, identify the major categories of cognitive processes by focusing on their core properties (Halford et al. 2014). Construction of representations in which elements are assigned to slots in a structure, and consequent ability to form structure-consistent mappings between representations, lie at the core of higher cognition. These abstract mappings are neither required nor evident in perception.

We are slicing nature at a joint by acknowledging the distinctiveness of major categories of cognition. To say this is not to deny the significance of processes that are common to many or all cognitions, but it entails acceptance of processes that are distinct to specific functions. Forming structured representations in working memory is a core process that is characteristic of higher cognition. Acceptance of these distinctions leads to a clarity that contributes to the conceptual coherence of the discipline.

Hallucinations and mental imagery demonstrate top-down effects on visual perception

doi:10.1017/S0140525X15002502, e248

Piers D. L. Howe and Olivia L. Carter
University of Melbourne, School of Psychological Sciences, The University of Melbourne, Victoria 3010, Australia.
pdhowe@unimelb.edu.au **ocarter@unimelb.edu.au**
http://psych.unimelb.edu.au/people/piers-howe
http://psych.unimelb.edu.au/people/olivia-carter

Abstract: In this commentary, we present two examples where perception is not only influenced by, but also in fact driven by, top-down effects: hallucinations and mental imagery. Crucially, both examples avoid all six of the potential confounds that Firestone & Scholl (F&S) raised as arguments against previous studies claiming to demonstrate the influence of top-down effects on perception.

In the target article, Firestone & Scholl (F&S) make the bold claim that higher-level cognition does not affect perception. They acknowledge that there have been a very large number of experiments that would seem to contradict them but argue that all of that previous work fell foul of at least one of six potential confounds. They challenged the academic community to find evidence for the effects of top-down cognition on perception that does not suffer from any of these potential confounds.

Hallucinations are one example that clearly meets the challenge. In psychotic disorders such as schizophrenia, hallucinations are most commonly reported in the auditory domain (e.g., hearing voices). However, pure visual hallucinations are also possible and are associated with hyperconnectivity between the amygdala and the visual cortex (Ford et al. 2015), underscoring the role of top-down feedback in their formation. There is also a range of natural and artificial compounds known to cause visual hallucinations when consumed. These drugs commonly induce vivid, geometric, kaleidoscope-type patterns behind closed eyes. Depending on the specific drug and dose, they can also lead to pure hallucinations involving creatures and scenes (Vollenweider 2001). Such drug-induced hallucinations are further demonstrations of top-down effects on perception. They are impossible to explain or conceptualise in terms of changes occurring within the bottom-up flow of external sensory information through the visual hierarchy.

Although it could be argued that we should consider such hallucinations an anomaly, independent of normal cognitive processes, other forms of hallucination are harder to dismiss. About 10%–30% of individuals who suffer from a severe visual impairment, such as glaucoma, experience Charles Bonnet syndrome (CBS; Schultz et al. 1996; Vukicevic & Fitzmaurice 2008). These individuals typically have no other neurological or psychiatric conditions, yet they frequently experience hallucinations. It is not that CBS sufferers merely think or sense the hallucinations; rather, the hallucinations appear so realistic that the affected people sometimes mistake them for reality (Schultz et al. 1996). In addition, the hallucinations will often interact with the participant's visual perception of the external world. For example, a hallucinated figure may obscure part of the visual scene that would otherwise be visible to the observer, thereby preventing the observer from seeing it – a clear example of top-down cognition influencing stimulus-based visual perception.

Mental imagery (i.e., visualisation) is a further example of top-down cognition obviously affecting perception. For some people, their mental images are exceptionally vivid, almost as vivid as visual perception (Pearson et al. 2011). Like hallucinations, these mental representations are pictorial: People actually see the mental images, as opposed to merely being aware of them (Kosslyn et al. 1995).

Both hallucinations and mental imagery activate the visual cortex in a way very similar to that of normal visual perception, which helps explain why people describe hallucinations and mental imagery as true perceptual experiences. Whereas hallucinations typically activate higher cortical areas (ffytche et al. 1998; Vollenweider 2001), mental imagery is able to additionally activate the primary visual cortex (Albers et al. 2013; Slotnick et al. 2005; Stokes et al. 2009). The BOLD (blood-oxygen-level dependent) fMRI activation caused by mental imagery is so similar to that generated by observed visual stimuli that a model based on tuning to low-level visual features (e.g., spatial frequency and orientation) that was trained on the BOLD fMRI activity generated when participants observed real images was able to determine which of those images the participants were subsequently visualising solely from the BOLD fMRI activity generated during these visualisations (Naselaris et al. 2015). Further, this low-level coupling between mental imagery and visual perception is unlikely to be epiphenomenal; magnetic pulses delivered to the primary visual cortex disrupted both mental imagery and visual perception to a similar extent (Blasdel & Salama 1986).

Hallucinations and mental imagery demonstrate that top-down cognitive and emotional processes can affect perception in a manner that avoids all six of the potential confounds F&S raise. There is no doubt that perception can be altered depending on which locations, features, or objects the observer attends to in the external world (Collins & Olson 2014; Vetter & Newen 2014). The fact that hallucinations and visual imagery are not driven by bottom-up stimuli, and that they are often unconstrained to specific locations in visual space, makes it difficult to conceptualize how such phenomena could be explained by such peripheral attention effects (confound 5) or indeed by low-level differences in the visual input (confound 4). Furthermore, because hallucinations and mental imagery are clearly perceptual, they also avoid confound 2 (perception vs. judgment) and confound 6 (memory and recognition). Additionally, they avoid the El Greco fallacy (confound 1) and cannot be attributed to demand and response bias (confound 3). As such, hallucinations and visual imagery avoid all of F&S's potential confounds.

In conclusion, it is clear that top-down processes can affect perception in a variety of ways. In clinical or drug-induced psychosis, a person's visual experience can be generated independent of bottom-up stimulation. In the case of mental imagery, percepts can be generated by directing top-down attention to internal mental states and representations. Finally, although CBS hallucinations appear to be beyond voluntary cognitive control (Schultz & Melzack 1991), they are clearly also caused by top-down processes. Moreover, such hallucinations can interact with visual perception, thereby providing a clear demonstration of a top-down effect influencing stimulus-based visual perception. These examples show that top-down cognitive processes are not only able to penetrate visual perceptions, but also they can cause them.

The distinction between perception and judgment, if there is one, is not clear and intuitive

doi:10.1017/S0140525X15002769, e249

Andreas Keller
Laboratory of Neurogenetics and Behavior, The Rockefeller University, New York, NY 10065.
andreasbkeller@gmail.com

Abstract: Firestone & Scholl (F&S) consider the distinction between judgment and perception to be clear and intuitive. Their intuition is based on considerations about visual perception. That such a distinction is clear, or even existent, is less obvious in nonvisual modalities. Failing to distinguish between perception and judgment is therefore not a flaw in investigating top-down effects of cognition on perception, as the authors suggest. Instead, it is the result of considering the variety of human perception.

The title of Firestone & Scholl's (F&S's) article, "Cognition does not affect perception: Evaluating the evidence for 'top-down' effects," suggests an article that investigates the effect of cognition on perception. The first sentence of the abstract ("What determines what we see?") reveals that the article is instead much more focused and discusses the effect of cognition on only a single modality: vision. Vision is of course the most important and interesting of the human modalities, and it is understandable that many researchers have specialized in its study. However, drawing conclusions about perception in general based on a single modality can be problematic. I will here illustrate this issue by discussing F&S's "Pitfall 2: Perception vs. Judgment," F&S call the distinction between perception and judgment "clear," (sect. 4.2, para. 1) "intuitive," and "uncontroversial (sect. 4.2.3)." Their evidence that the distinction is uncontroversial: Perception and judgment can be in clear conflict in visual illusions.

However, the intuition that there is a clear and uncontroversial distinction between perception and judgment does not generalize well to other modalities. Take, for example, the question whether we perceive or judge that a shoe is uncomfortable. The authors' proposal is that we can judge, but not perceive, whether a shoe is comfortable or not. That is intuitively true when *perceiving* means *seeing*. We can look at the shoe and perceive that it is smaller and of a different shape than we remember our own foot to be. From that, we can judge that the shoe is uncomfortable. That is a judgment and not a perception because, in addition to the perception of the size of the shoe (and the memory of the perceived size of our foot), we also need to perform a mental comparison of the two sizes to arrive at the conclusion that the shoe is uncomfortable. When we broaden the meaning of *perceiving* to include all sensory modalities, it is much less obvious whether the uncomfortableness of the shoe is judged or perceived. If wearing the shoe hurts, do we perceive that it is an uncomfortable shoe, or should that also be considered a judgment because it involves an extra step from perceiving the pain to judging the shoe to be uncomfortable? If the left shoe hurts more than the right shoe, do we perceive the difference, or is anything that involves a comparison a judgment? The answers to such questions are not clear and intuitive.

The example of the uncomfortable shoe shows that outside of vision, the distinction between perception and judgment is not uncontroversial. The prime example of a modality in which it is impossible to disentangle the two is olfaction. Plato wrote that odors "have no name and they have not many, or definite and simple kinds; but they are distinguished only as painful and pleasant" (Timaeus 67a). More recently, multidimensional scaling techniques confirmed that valence is the most important perceptual dimension in olfaction (Haddad et al. 2008). In humans, olfaction has evolved to be an evaluative sense. Olfactory information is used mainly to make decisions about rejecting or accepting food, mates, or locations (Stevenson 2009). Put differently, more "than any other sensory modality, olfaction is like emotion in attributing positive (appetitive) or negative (aversive) valence to the environment" (Soudry et al. 2011, p. 21). Olfaction is a judgmental sense in which perceiving and judging are intertwined.

In summary, contrary to what F&S write, authors talking about "perceptual judgment" (sect. 4.2.3, para. 2) do not invite confusion about a foundational distinction between perception and judgment. Instead, they present evidence that there is no foundational distinction between perception and judgment. Consequently, the failure to disentangle perception from judgment is not a pitfall of flawed studies, but rather an acknowledgment that there is no clear division between the two. Although such a claim may seem revolutionary for vision, it is not a new idea for other modalities. We should not make the mistake of basing our understanding of perception exclusively on vision.

Cognition can affect perception: Restating the evidence of a top-down effect

doi:10.1017/S0140525X15002642, e250

Daniel T. Levin,[a] Lewis J. Baker,[a] and Mahzarin R. Banaji[b]

[a]Department of Psychology and Human Development, Vanderbilt University, Nashville, TN 37203-5721; [b]Department of Psychology, Harvard University, Cambridge, MA 02138.

daniel.t.levin@vanderbilt.edu **lewis.j.baker@vanderbilt.edu**
mahzarin_banaji@harvard.edu
https://my.vanderbilt.edu/daniellevinlab/about-me/
https://my.vanderbilt.edu/lewisbaker/about-me/
http://www.people.fas.harvard.edu/~banaji/

Abstract: We argue that Firestone & Scholl (F&S) provide worthwhile recommendations but that their critique of research by Levin and Banaji (2006) is unfounded. In addition, we argue that F&S apply unjustified level of skepticism about top-down effects relative to other broad hypotheses about the sources of perceptual intelligence.

We believe that Firestone & Scholl's (F&S's) target article represents a commendable standard for evaluating top-down effects, and we agree with many of their recommendations. We also agree that it is possible to overstate the power of fleeting abstractions to impact our immediate impression of the world, and we are skeptical of the view that perception is so saturated with belief that we can see whatever we wish. However, we also know from many decades of research that perception integrates sensory input with reliable world-knowledge. To deny such evidence would be to deny that humans are flexible learners. This is where we probably diverge from F&S: We do not think that evidence for top-down processes represents a surprising or dramatic departure from established theory, nor that the top-down hypothesis requires a higher standard of proof than any other hypothesis about the functioning of the mind in physical and social space.

Here, we focus on the impact of race on the perception of the lightness of faces. Levin and Banaji (2006) demonstrated that participants seem to perceive Black faces to be darker than White faces. This effect was present both when participants adjusted samples to match unambiguous Black and White faces, and when one group of participants judged an ambiguous face that they were told was Black while another group saw the same face, this time believing that it was White. In the target article, F&S review previous experiments (Firestone & Scholl 2015a) that focus on the former effect with unambiguous faces, arguing that it was a qualitatively more effective demonstration of a top-down effect. However, they also argued that this finding suffered a potential stimulus confound. F&S tested this confound using blurred faces (the left half of Fig. 1) to measure whether participants who were nominally unable to identify the faces by race still showed the lightness illusion (e.g., participants who indicated that the faces were of the same race still judged the White face to be lighter).

In our response (Baker and Levin 2016) we noted that F&S used a forced-choice question to obtain judgments about which face was darker, and used a nonforced choice to assess detection of race (participants selected from a menu of possible races independently for each face). What if the forced-choice lightness question was more sensitive to lightness than the race-detection question was to race? F&S may have underestimated participants' ability to detect race in the blurred faces, and may therefore have falsely classified some participants as unable to detect race. This seems particularly plausible given that the classification was based on one or two judgments about subtle, near-threshold information. Indeed, when we included a forced-choice question that directly asked participants to choose which face was White and which was Black, we repeatedly observed that 75%–80% of participants correctly assigned race. It is important to note that participants were just as successful in detecting the race of the faces when the faces were contrast-inverted (Fig. 1), so it seems unlikely that they detected the race of the faces by noting the brightness confound and by guessing that the lighter face was White.

In addition to evidence that the blurring left some race-specifying information in the images, Baker and Levin found that participants who correctly distinguished the race of the noninverted faces also were more likely to judge that the Black face was darker. This result supports the hypothesis that there is a relationship between participants' ability to

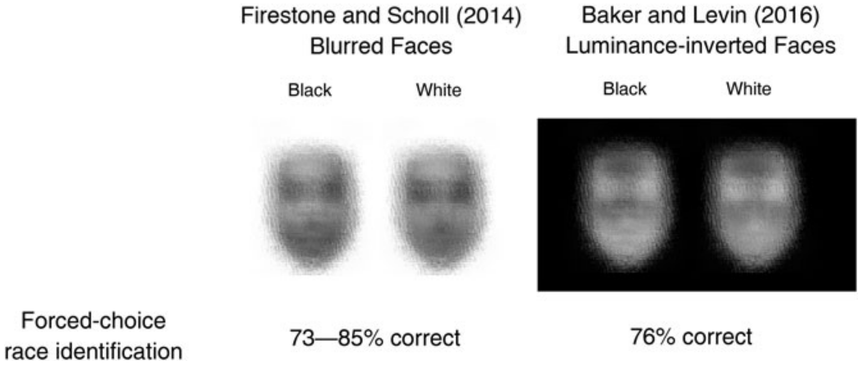


Figure 1 (Levin et al.). Illustration of accuracy on forced-choice identification of race for blurred stimuli employed by Firestone and Scholl (2015a). Firestone and Scholl argued that participants who could not identify the race of the stimuli on left nonetheless showed a brightness effect. However, we demonstrated that participants were able to accurately judge the race of the faces both for the original blurs and in luminance-inverted versions of the faces (Baker and Levin 2016).

perceive the race of the faces and how light each face appears to them.

Space constraints prevent us from reviewing all of F&S's critique and all of the logic underlying our response, but the complexity of the issue leads us to a key point: The original Levin and Banaji report foresaw the difficulty in fully eliminating confounds inherent to two different stimuli, and so it included the abovementioned ambiguous face experiment, along with an experiment in which RT on same–different judgments was slowed when a relatively lightened Black face was compared with a White face (thus equalizing their apparent lightnesses). These additional experiments cast serious doubt on F&S's conclusion "that the initial demonstration of Levin and Banaji (2006) provides *no evidence* for a top-down effect on perception" (sect 4.4.1, para. 5; emphasis added). The casual reader might be forgiven for assuming that Levin and Banaji's entire study can be dismissed unless they realize that the word "initial" means that only one of several experiments are at issue and read the footnote describing one of these other experiments. We think that this quote reveals a fundamental problem with F&S's approach. The categorical conclusion implies that experiments must either provide unambiguous proof of top-down effects by avoiding all of the pitfalls they describe, or the work falls to zero weight in tipping the scale to the top-down side of a debate that is complex enough to have been raging for a long time.

We prefer a more nuanced approach to advancing research on this topic for several reasons. First, there are many different kinds of top-down effects, some in which momentary thoughts influence how things look, and some more subtle effects where a more-sophisticated perceptual process influences a less-sophisticated one, perhaps as the result of long-term experience. This is especially evident in the social domain, where category-informed reactions to skin color can clearly be consequential. Of course, researchers' specific interests might lead them to isolate the truly perceptual sources of judgments about experience, but at some point it becomes an exercise in purity that provides license to focus exclusively on relatively artificial stimuli and tasks designed a priori to reveal phenomena that will confirm evidence of bottom-up processing. In all cases rigor is crucial, and F&S provide some good recommendations in achieving that. But rigor should not be an excuse to ignore the study of important phenomena. We believe that discovery is best served by exploring the full richness of human perceptual capacities that may or may not reveal cognitive penetration rather than dwelling exclusively on simpler perceptual process from a penchant for tidiness.

Not even wrong: The "it's just X" fallacy

doi:10.1017/S0140525X15002721, e251

Gary Lupyan
Department of Psychology, University of Wisconsin–Madison, Madison, WI 53706.
lupyan@wisc.edu
http://sapir.psych.wisc.edu

Abstract: I applaud Firestone & Scholl (F&S) in calling for more rigor. But, although F&S are correct that some published work on top-down effects suffers from confounds, their sweeping claim that there are no top-down effects on perception is premised on incorrect assumptions. F&S's thesis is wrong. Perception is richly and interestingly influenced by cognition.

Disagreements arise when people argue with different facts. But disagreements can also arise when people argue from different starting assumptions. F&S and I share all of the same facts, but F&S come to the wrong conclusions because they have the wrong assumptions.

Many of the studies F&S review indeed suffer from stimulus and experimenter-demand confounds, But many others are well-controlled investigations using gold-standard psychophysical methods. These studies show that expectations and knowledge affect virtually all aspects of visual perception. For example, knowledge of surface hardness affects amodal completion (Vrins et al. 2009), knowledge of bodies affects perceiving depth from binocular disparity (Bulthoff et al. 1998), expectations of motion affect motion perception (Sterzer et al. 2008), and knowledge of real-world size affects perceived speed of motion (Martín et al. 2015). Meaningfulness – a putatively late process – affects putatively earlier processes such as shape discrimination (Lupyan & Spivey 2008; Lupyan et al. 2010) and recovery of 3-D volumes from two-dimensional images (Moore & Cavanagh 1998). Color knowledge affects color appearance of images (Hansen et al. 2006) and even color afterimages (Lupyan 2015b). Hearing a word affects the earliest stages of visual processing (Boutonnet & Lupyan 2015; see also Landau et al. 2010; Pelekanos & Moutoussis 2011).

How can F&S, who are aware of all of this work (some of which they discuss in detail in the target article), still argue that there are no top-down effects on perception? They dismiss all of those studies on the grounds that they are "just" effects of attention, memory, or categorization/recognition. This "it's not perception, it's just X" reasoning assumes that attention, memory, and so forth be cleanly split from *perception proper*. But attentional effects can be dismissed if and only if attention simply changes input to a putatively modular visual system (sect. 4.5). Memory effects can be dismissed if and only if memory is truly an amodal "back-end" system. Recognition and categorization effects can be dismissed if and only if these processes are wholly downstream of "true" perception (sects. 3.4, 4.6). All of those assumptions are wrong.

Some aspects of attention really *are* a bit like changing the input to our eyes. Attending to one or another part of a Necker cube is kind of like shifting one's eyes. If we dismiss the latter as an interesting sort of top-down effect on perception, we should likewise dismiss the former. But as we now know, attention is far richer. We can, for example, attend to people or dogs, or the letter "T" (across the visual field) – a process of deploying complex priors within which incoming information is processed. In so doing, attention warps the visual representations (e.g., Çukur et al. 2013; sect. 5.2 in Lupyan 2015a for discussion). Aside from the simplest confounds in spatial attention, attentional effects are not an alternative to top-down effects on perception, but rather one of the mechanisms by which higher-level knowledge affects lower-level perceptual processes (Lupyan & Clark 2015).

Some top-down effects can be dismissed as being effects on memory. Someone might remember a $20 bill as being larger than a $1 bill, but not see it as such. But F&S's "just memory" argument goes much further. For example, Lupyan and Spivey (2008) found that instructing participants to view the meaningless symbols ⊓⊔ and ⊓⊔ as meaningful – rotated numbers 2 and 5 – improved visual search efficiency. F&S argue that this might be merely an effect on memory, citing Klemfuss et al. (2012) as having shown that decreasing the memory load by showing participants a target-preview caused the meaningfulness advantage to disappear. But actually, the largest effect of the target-preview was to *slow* search performance for the meaningful-number condition, bringing it in line with that of the meaningless-shape condition.

But suppose Klemfuss et al. actually found that showing a target-preview to participants improved search as much as the instructional manipulation we had used. Would this mean that meaningfulness does not affect perception? Not at all! If telling people to think of ⊓⊔ and ⊓⊔ as 2s and 5s is as effective as showing a target preview in helping them to find the completely unambiguous target in a singleton search, that would mean a high-level instructional manipulation meaning can affect visual search efficiency as much as an overtly visual aid. A top-down effect that can be partially ascribed to memory does not mean it is not (also) an effect on perception, because part of what we call memory – visual memory – appears to have a perceptual locus (D'Esposito & Postle 2015; Pratte & Tong 2014). This is why holding visual items in memory causes people to see things differently (e.g., Scocchia et al. 2013).

Visual memory is not a back-end system, as F&S assume. It is perceptual. This helps explain the confusion F&S have about Lupyan and Ward's (2013) demonstration that hearing a word (e.g., "kangaroo") can make visible an image of a kangaroo made invisible through continuous flash suppression. Lupyan & Ward's explanation was exactly the same as F&S's (sect. 4.6.2): Hearing a word activates visual knowledge – knowledge that is *visual* – which we argued allows people to see otherwise weak and fragmented visual inputs. Even if this is "merely" an effect on back-end memory, the fact remains that hearing a word improves sensitivity in simply detecting objects. It helps people *see*. Like attention, memory is part of the mechanism by which knowledge affects perception.

Lastly, recognition. Vision scientists might be surprised to learn that, according to F&S, studying how people recognize a dog as a dog or that two objects look the same is studying the postperceptual back end, but studying animacy (Gao et al. 2009), causal history (Chen & Scholl 2016), and reconstructions of shapes through occluders (Firestone & Scholl 2014a), is studying true perception. Not all perceptual tasks require recognition, but most of the ones vision scientists care about do. If simply detecting an object as an object (Lupyan & Ward 2013) is "just" recognition and therefore not true perception, many vision scientists might want to find other employment.

Representation of affect in sensory cortex

doi:10.1017/S0140525X15002708, e252

Vladimir Miskovic,[a] Karl Kuntzelman,[a] Junichi Chikazoe,[b,c] and Adam K. Anderson[b]

[a]*Department of Psychology, State University of New York at Binghamton, Binghamton, NY 13902;* [b]*Department of Human Development, Human Neuroscience Institute, Cornell University, Ithaca, NY 14853;* [c]*Section of Brain Function Information, National Institute for Physiological Sciences, Okazaki, Aichi, 444-8585, Japan*

miskovic@binghamton.edu **kkuntze1@binghamton.edu**
chikazoe@nips.ac.jp **aka47@cornell.edu**

Abstract: Contemporary neuroscience suggests that perception is perhaps best understood as a dynamically iterative process that does not honor cleanly segregated "bottom-up" or "top-down" streams. We argue that there is substantial empirical support for the idea that affective influences infiltrate the earliest reaches of sensory processing and even that primitive internal affective dimensions (e.g., goodness-to-badness) are represented alongside physical dimensions of the external world.

Although we believe that Firestone & Scholl (F&S) offer sound advice for culling theoretical excesses, here we argue that (1) contemporary advances within the neurosciences constitute legitimate challenges for foundational concepts invoked within the classical cognitive architecture of perception, and (2) the wider literature in fact does provide compelling support for the effects of motivational factors on perceptual processes.

The commentary authors briefly consider how systems neuroscience might inform their discussion, but they largely dismiss what knowledge of "descending neural pathways" (sect. 2.2) might be able to contribute to this debate on the basis that such knowledge is not novel. We respectfully disagree with that position. It must eventually be possible to relate any useful cognitive architecture to neurophysiology, and hence it is desirable to respect such constraints when they are known and interpretable. Franconeri et al. (2013), for example, used well-established principles of cortical organization to provide substance to the previously fuzzy concept of cognitive "resources" in a way that is intuitive and coherent.

Let us take visual cortex (V1) as the prototypical sensory system. Classical feedforward feature-detector models account for only about 40% of the variance in V1 function (Carandini et al. 2005). In a slightly more pessimistic estimate, Olshausen and Field (2005) quantify our epistemic uncertainty concerning V1 functional properties to be closer to 85%. This point is not meant to disparage the remarkable progress of visual neuroscience research, but simply to demonstrate that it seems premature to claim an understanding of perception's architecture that is comprehensive enough to preclude original neurophysiological insights. Our current state of uncertainty about the earliest stages of vision appears less surprising if we consider a few salient neuroanatomical facts. Only about 5% of the excitatory synaptic input to layer IV of V1 derives from geniculate drive (Douglas & Martin 2007), and approximately 60%–80% of V1 responses are attributable to other V1 neurons or nongeniculate inputs (Muckli & Petro 2013). Such structural features ensure that primary visual cortex sustains multiple interactions with high-level sources of information. Recent developments in high-resolution tract tracing (Markov & Kennedy 2013) and considerations of electrophysiological timing (Briggs & Usrey 2005) suggest that perception is more properly conceptualized as a dynamically reverberating loop rather than encapsulated bottom-up and top-down streams. This revised conceptualization makes the prospect of carving the underlying machinery at the joints much more daunting. Indeed, within contemporary systems neuroscience, the organism's internal context is recognized to be as important for unraveling the nature of sensory processing as are the physical parameters of stimuli (Fontanini & Katz 2008). From a neurobiological perspective, it is not so much a question of *if* perception is penetrable, but to what degree and under what circumstances the dynamics might change (Muckli 2010).

Research conducted in alert animals has long recognized that motivational factors rapidly and profoundly influence neuronal responses at the earliest modality-specific perceptual stages, resulting in increased gain and/or altered tuning curves (McGann 2015). Findings collected across a range of mammalian species have demonstrated tonotopic map remodeling within primary auditory cortex that optimizes the processing of tones paired with rewards or punishers (Weinberger 2004). The amount of representational area expansion for conditioned stimuli in primary sensory cortices may encode the magnitude of affective relevance (Rutkowski & Weinberger 2005) and even predict subsequent extinction learning (Bieszczad & Weinberger 2010). Research in humans has also documented pronounced time-varying changes for conditioned cues across several sensory cortices (Miskovic & Keil 2012).

The preceding cases are well-established but modest demonstrations of perceptual penetrability. We would like to advance as a hypothesis a strong version of penetrability, according to which primitive affective qualities such as hedonic valence might be understood as perceptual attributes that are represented alongside other, more objective, physical properties. This strong version is therefore closer in nature to Wündt's (1897) insights about the central role of affect in perception. This proposition suggests that the affective dimensions of perceptual experience enjoy a neural currency that is not altogether dissimilar from the dimensions that reflect the physics of stimuli (e.g., light wavelength).

Elementary valence attributes might be embedded within modality-specific sensory cortices in population codes – distributed activity, within or across brain regions, that represents the relationships between stimulus or experiential properties and their distances in a high-dimensional space (Kriegeskorte & Kievit 2013). We recently employed a representational similarity analysis of blood-oxygen-level dependent (BOLD) signals to examine how external events are represented as pleasant or unpleasant, alongside other physical (e.g., low or high luminance) and semantic (e.g., representing either animate or inanimate objects) properties (Chikazoe et al., 2014). We found that activity patterns in the ventral temporal cortex and the anterior insular cortex contained the representational geometry of modally bound valence representations belonging to the visual and gustatory systems, respectively. In addition to such modality-specific representations, we also found evidence for a population code in orbitofrontal cortex that is shared across events originating from distinct modalities, which presumably allows subjective affect to be objectively quantified and compared on a common valence axis. That an aversive image and an acrid taste are both experienced as hedonically unpleasant may therefore be by virtue of their objective similarity in distributed neural population codes – a transfer function is interposed between purely physical sensations and their elementary valence. Whereas 680 nm of light might carry information related to the perceptual experience of red, how the affective tone of experience affects the observer may emerge from higher-level processes that become partially embedded within perceptual representations.

In short, we believe that analogies to dynamically reverberating loops and principles of reciprocal causation provide a much closer approximation to the ways that brains function, and that these ideas necessitate a more thoroughgoing reevaluation of many cognitivist axioms. It is quite possible – indeed it seems likely – that static distinctions between perception, cognition, and emotion reflect much more about historical intellectual biases in the field of cognitive science than about the true operations of the brain/mind.

Beyond perceptual judgment: Categorization and emotion shape what we see

doi:10.1017/S0140525X15002514, e253

Steven B. Most
School of Psychology, UNSW, Australia, NSW 2052, Australia.
s.most@unsw.edu.au

Abstract: By limiting their review largely to studies measuring perceptual judgment, Firestone & Scholl (F&S) overstate their case. Evidence from inattentional blindness and emotion-induced blindness suggests that categorization and emotion shape what we perceive in the first place, not just the qualities that we judge them to have. The role of attention in such cases is not easily dismissed as "peripheral."

Firestone & Scholl (F&S) parry recently resurgent suggestions that higher-order aspects of the mind – such as motivation, emotion, and categorization – alter perception, and they highlight several pitfalls endemic to such claims. Their paper is a valuable contribution to the literature.

However, in staking out territory beyond the bounds of early vision, F&S make the overly sweeping claim that no studies have provided evidence for top-down effects on "*what we see* as a whole visual processing and the conscious percepts it produces" (sect. 1.2, para. 2; emphasis theirs). A sweeping claim requires a sweeping survey of the literature, but unfortunately the authors largely limit their review to studies claiming that higher-order factors make things look closer, bigger, steeper, darker, or wider. In other words, they focus largely on studies that measure perceptual judgments.

What of studies that measure whether people see a stimulus at all? The inattentional blindness and emotion-induced blindness literatures respectively suggest that categorization and emotion mold our ability to see things in the first place Although the authors appear to dismiss such phenomena as "peripheral" effects of attention (sect. 4.5), attention guides perception at multiple points in visual processing, and it is likely incorrect to relegate its role only to the gating of input into early vision. (The authors vaguely concede that attention is not always so peripheral, but ultimately, they gloss over this point.) Evidence indeed suggests that these effects cannot be waved away as merely a result of peripheral selection (nor to memory, which the authors identify as another potential pitfall).

Inattentional blindness. The ability to see that something is present at all depends on more than where we direct our eyes. Nowhere is this more apparent than in the phenomenon of inattentional blindness: the failure to see obvious and salient stimuli that people look directly at while their attention is preoccupied (Mack & Rock 1998; Most 2010; Most et al. 2001; 2005a; 2005b; Simons & Chabris 1999).

Inattentional blindness experiments suggest that categorization shapes perception in ways that cannot be attributed to selection for visual input. In one study, participants either tracked four moving digits and ignored four moving letters or tracked the letters while ignoring the digits (Most 2013). When an unexpected *E* traveled across the screen, those tracking the letters were more likely to notice it than those tracking the digits, and this pattern reversed when the unexpected object was its mirror image, a block-letter *3*. Because this effect was driven by the categorization of the unexpected object, it must have stemmed from selection after a degree of visual processing had already occurred. Evidence further suggests that the role of categorization in inattentional blindness is unlikely caused by confusions between the unexpected object and other members of the nontarget set (Koivisto & Revonsuo 2007). Nor do inattentional blindness effects appear to be wholly attributable to failures of memory (Ward & Scholl 2015).

Emotion-induced blindness. In *emotion-induced blindness (EIB)*, people view rapid serial visual presentations of items and search for a single target within each stream. When the target is preceded by an emotionally powerful picture, people are unable to report the target (Most et al. 2005a).

At first glance, EIB seems like something that could fall within one of two categories of F&S's pitfalls. First, because emotional stimuli capture attention (MacLeod et al. 1986), attention could simply be too preoccupied to guide input of the target into the visual system (sect. 4.5). Second, it could be that EIB is more a phenomenon of memory than of perception (sect. 4.6), either because the measure is typically retrospective (people are usually asked to report the target at least half a second after its appearance; see Wolfe 1999) or because it seems phenomenally related to the attentional blink, which itself has been attributed to failures to consolidate information into visual working memory (e.g., Chun & Potter 1995). However, there are reasons to suspect that neither peripheral aspects of attention nor memory are accountable.

Specifically, in contrast to studies demonstrating that emotional stimuli capture attention to their location, in EIB target perception is worse *at* the location of the emotional distractor (e.g., Most & Wang 2011). This pattern has led to suggestions that emotional distractors compete for neural representation with targets that appear in the same receptive field (Wang et al. 2012; also see Keysers & Perrett 2002), a suggestion consistent with findings that neural responses to targets and emotional distractors exhibit a trading relationship (Kennedy et al. 2014). Because people tend to prioritize emotional information, it is the emotional distractors that win, an effect that seems to be modulated by mood (Most et al. 2010).

As with inattentional blindness, EIB appears to be a perceptual phenomenon rather than a memorial one. When participants were instructed to respond to targets immediately upon seeing them rather than at the end of each trial, the effect was undiminished (Kennedy & Most 2012). In addition, the spatially localized nature of the effect suggests that it arises from competition at a stage prior to consolidation into working memory.

In sum, F&S provide an incisive critique of the literature on top-down effects on perception. But they overstate their case when claiming that evidence for such effects is nonexistent. Surveying the literature on *what* people do or don't see rather than on their perceptual judgments reveals instances that don't easily fall within the categories of pitfall F&S outline. Of course, the authors may argue that such effects still fall outside the bounds of "perception," strictly defined, and that they fall prey to other misconceptions about the relationship between perception and cognition. So far, however, while stating that perception extends beyond the computations involved in "early vision," they have left the placement of the line between perception and cognition ambiguous. Good fences make good neighbors; as the target article's authors no doubt consider perception and cognition to be neighboring – if not encroaching – domains of the mind, it would be neighborly to provide a better map of where they think that fence stands.

Convergent evidence for top-down effects from the "predictive brain"[1]

doi:10.1017/S0140525X15002599, e254

Claire O'Callaghan,[a] Kestutis Kveraga,[b] James M. Shine,[c] Reginald B. Adams, Jr.,[d] and Moshe Bar[e]

[a]*Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge CB2 3EB, United Kingdom;* [b]*Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02129;* [c]*School of Psychology, Stanford University, Stanford, CA 94305;* [d]*Department of Psychology, The Pennsylvania State University, University Park, PA 16801;* [e]*Gonda Center for Brain Research, Bar-Ilan University, Ramat Gan 5290002, Israel.*

co365@cam.ac.uk **kestas@nmr.mgh.harvard.edu**
macshine@stanford.edu **radams@psu.edu** **moshe.bar@biu.ac.il**

Abstract: Modern conceptions of brain function consider the brain as a "predictive organ," where learned regularities about the world are utilised to facilitate perception of incoming sensory input. Critically, this process hinges on a role for cognitive penetrability. We review a mechanism to explain this process and expand our previous proposals of cognitive penetrability in visual recognition to social vision and visual hallucinations.

A neural mechanism for cognitive penetrability in visual perception. In their target article, Firestone & Scholl (F&S) readily dismiss the extensive presence of descending neural pathways (Angelucci et al. 2002; Bullier 2001), claiming they have no necessary implications for cognitive penetrability. Yet it is precisely this architecture of feedforward and feedback projections, beginning in the primary visual cortex, ascending through dorsal or ventral visual pathways, dominated respectively by magnocellular and parvocellular cells (Goodale & Milner 1992; Ungerleider & Mishkin 1982), and matched with reciprocal feedback connections (Felleman & Van Essen 1991; Salin & Bullier 1995), that provides the starting point for evidence in favour of top-down effects on visual perception.

Numerous studies capitalised on inherent differences in the speed and content of magnocellular (M) versus parvocellular (P) processing to reveal their role in top-down effects (Panichello et al. 2012). Early work using functional magnetic resonance imaging (fMRI) suggested that formation of top-down expectations, signalled by a gradually increasing ventrotemporal lobe activity, facilitated the recognition of previously unseen objects (Bar et al. 2001). In subsequent studies using both intact line drawings and achromatic, low spatial frequency (LSF) stimuli (which preferentially recruit M pathways), early activity was evident in the orbitofrontal cortex (OFC) ~130 ms after stimulus presentation, well before object recognition–related activity peaks in the ventrotemporal cortex (Bar et al. 2006). An fMRI study using dynamic causal modelling later confirmed that M-biased stimuli specifically activated a pathway from the occipital cortex to OFC, which then initiated top-down feedback to the fusiform gyrus. This connectivity pattern was different from that evoked by stimuli activating the P pathway, where only feedforward flow increased between the occipital cortex and fusiform gyrus (Kveraga et al. 2007a). OFC activity predicted recognition of M, but not P, stimuli, and resulted in faster recognition of M stimuli by ~100 ms. Another fMRI study showed that this OFC facilitation of object recognition was triggered for meaningful LSF images exclusively: Only meaningful images, but not meaningless images (from which predictions could not be generated), revealed increased functional connectivity between the lateral OFC and ventral visual pathway (Chaumon et al. 2013). We argue not only that these results demonstrate the importance of descending neural pathways, which F&S do not dispute (cf. sect. 2.2), but also that these recurrent connections penetrate bottom-up perception and facilitate perception via feedback of activated information from the OFC.

This top-down activity does not merely reflect recognition or "back-end" memory-based processes, as F&S suggest are commonly conflated with top-down effects. Instead, the rapid onset of OFC activation and subsequent coupling with visual cortical regions indicate that these top-down processes affect perception proper, which we suggest occurs in the form of predictions that constrain the ongoing perceptual process. These predictions categorise ambiguous visual input into a narrow set of most probable alternatives based on all available information. As a richly connected association region, receiving inputs from sensory, visceral and limbic modalities, the OFC is ideally situated to integrate crossmodal information and generate expectations based on previous experience that can be compared with incoming sensory input. Predictive information from the OFC is then back-propagated to inferior temporal regions and integrated with high spatial frequency information. Thus, by constraining the number of possible interpretations, the OFC provides a signal that guides continued, low-level visual processing, resulting in a refined visual percept that is identified faster (Trapp & Bar 2015).

Another aspect of this top-down guidance process that can penetrate bottom-up visual perceptual processing involves constraints imposed by the stimulus context. In the model that emerged from these data, "gist" information is extracted from LSFs in the visual input, and predictions are generated about the most probable interpretation of the input, given the current context (Bar 2004). When bottom-up visual input is ambiguous, the same object can be perceived as a hair dryer or a drill, depending on whether it appears in a bathroom or a workshop context (e.g., Bar 2004, Box 1). A network sensitive to contextual information, which also includes the parahippocampal, retrosplenial, and medial orbitofrontal cortices (Aminoff et al. 2007; Bar & Aminoff 2003), has been implicated in computing this context signal. Crucially, this process is not simply influencing better guesswork. Using magnetoencephalography and phase synchrony analyses, these top-down contextual influences are shown to occur during the formation stages of a visual percept, extending all the way back to early visual cortex (Kveraga et al. 2011).

The emerging picture from this work suggests that ongoing visual perception is directly and rapidly influenced by previously learnt information about the world. This is undoubtedly a highly adaptive mechanism, promoting more efficient processing amidst the barrage of complex visual input that our brains receive. In the next section, we extend this model by incorporating ecologically valid examples of how top-down effects on visual perception facilitate complex human interactions, and the ramifications when the delicate balance between prediction and sensory input is lost in clinical disorders.

Cognitive penetrability in broader contexts – visual hallucinations and social vision. Top-down influences on visual perception are also observable in clinical disorders that manifest visual hallucinations, including schizophrenia, psychosis, and Parkinson's disease. Most strikingly, the perceptual content of visual hallucinations can be determined by autobiographical memories; familiar people or animals are a common theme (Barnes & David 2001). Frequency and severity of visual hallucinations is exacerbated by mood and physiological states (e.g., stress, depression, and fatigue), with mood also playing an important role in determining the content of hallucinations (e.g., when the image of a deceased spouse is perceived during a period of bereavement) (Waters et al. 2014). Such phenomenological enquiry into visual hallucinations suggests their content is influenced in a top-down manner, by stored memories and current emotional state. Anecdotal report is mirrored by experimental confirmation that the psychosis spectrum is associated with overreliance on prior knowledge, or predictive processing, when interpreting ambiguous visual stimuli (Teufel et al. 2015). Together, these mechanisms are consistent with a framework in which top-down influences tend to dominate visual processing in hallucinations. Important for theories of cognitive penetrability, visual hallucinations typically involve a hallucinated object being perceived as embedded within the actual scenery, such that the hallucination is thoroughly integrated with sensory input (Macpherson 2015). Existing neural frameworks for visual hallucinations account for an imbalance between bottom-up sensory information and top-down signals. These frameworks implicate overactivity in regions supplying top-down information during normal visual perception, including the medial temporal and prefrontal sites in the model outlined above. Abnormal activity in these regions, and in their connectivity with the visual cortex, plays a causative role in creating the hallucinatory percepts that effectively hijack visual perception (Shine et al. 2014). Electrical stimulation studies targeting these regions independently confirm that abnormal activity in temporal lobe and midline areas is capable of generating complex visual hallucinations (Selimbeyoglu & Parvizi 2010).

Visual information conveying social cues is some of the subtlest, yet richest, perceptual input we receive – consider the abundance of information delivered in a sidelong glance or a furrowed brow. Top-down influences allowing us to recognise patterns in our social environment and interpret them rapidly are a cornerstone of adaptive social behaviour (de Gelder & Tamietto 2011). Available evidence suggests that social visual processing leverages precisely the same neural mechanism described above for object and scene recognition. However, because of typically greater ambiguity in social cues, social vision must rely on top-down expectations to an even greater extent than object recognition. Social cues, including eye gaze, gender, culture, and race are found to directly influence the perception and neural response to facial emotion (Adams & Kleck 2005; Adams et al. 2003; 2015) and exert increasing influence with increasing ambiguity in the given expression (Graham & LaBar 2012). Critically, these effects are also modulated by individual differences such as trait anxiety and progesterone levels in perceptual tasks (Conway et al. 2007; Fox et al. 2007) and in amygdala response to threat cues (Ewbank et al. 2010). Dovetailing with the model outlined above, fusiform cortex activation is found to track closely with objective gradations between morphed male and female faces, whereas OFC responses track with categorical perceptions of face gender (Freeman et al. 2010). As in object recognition, OFC may be categorising continuously varying social stimuli into a limited set of alternative interpretations. Social vision therefore provides an important example of cognitive penetrability in visual perception that utilises stored memories and innate templates to makes sense of perceptual input.

Conclusion. The convergent experimental and ecological evidence we have outlined suggests a visual processing system profoundly influenced by top-down effects. The model we describe fits with a "predictive brain" harnessing previous experience to hone sensory perception. In the face of evidence reviewed here, it seems difficult to categorically argue that cognitive penetrability in visual perception is yet to be convincingly demonstrated.

NOTE

1. Claire O'Callaghan and Kestutis Kveraga are co-first authors of this commentary.

Firestone & Scholl conflate two distinct issues

doi:10.1017/S0140525X15002459, e255

Ryan Ogilvie and Peter Carruthers
Department of Philosophy, University of Maryland, College Park, MD 20742-7615.
rogilvie@umd.edu pcarruth@umd.edu
https://sites.google.com/site/ryanogilvie/
http://faculty.philosophy.umd.edu/pcarruthers/

Abstract: Firestone & Scholl (F&S) seem to believe that the viability of a distinction between perception and cognition depends on perception being encapsulated from top-down information. We criticize this assumption and argue that top-down effects can leave the distinction between perception and cognition fully intact. Individuating the visual system is one thing; the question of encapsulation is quite another.

What is at stake in the debate between those who think that vision is informationally encapsulated and those who don't? Firestone & Scholl (F&S) believe that it is the viability of the distinction between perception and cognition. They write, "the extent to which what and how we *see* is functionally independent from what and how we think, know desire, act, and so forth" bears on "whether there is a salient 'joint' between perception and cognition" (sect. 2). They further claim that if cognition "can affect what we see…then a genuine revolution in our understanding of perception is in order" (sect. 1.1, para. 3). Thus, they seem to believe that one can draw the traditional distinction between perception and cognition *only if* perception is encapsulated from cognition.

Although F&S cite a number of opposing authors who agree with this sentiment, we think that it rests on a conflation between two different notions of modularity. (Roughly, the distinction is between Fodor modularity, which requires encapsulation [see Fodor 1983] and functional modularity, which does not [see Barrett & Kurzban 2006; Carruthers 2006].) In fact, one can perfectly well individuate a system in terms of what it does (e.g., what computations it performs) *regardless of whether* the operations of the system are sensitive to information from outside.

One can characterize the visual system as the set of brain mechanisms specialized for the analysis of signals originating from the retina. The computations these mechanisms perform are geared toward making sense out of the light array landing on the retina. We may not know precisely how to identify these mechanisms or how they perform their computations. But it is surely a plausible hypothesis that there *is* a set of brain mechanisms that does this, and perhaps only this. (Indeed, it is widely assumed in the field that the set would include at least V1, V2, and V3.) In this we agree with F&S. But one can accept that the visual system consists of a proprietary set of mechanisms while denying that it takes only bottom up input. For example, the existence of crossmodal effects need in no way undermine the distinction between audition and vision. Hence, we see no reason to think that the existence of top-down effects should undermine the distinction between vision and higher-level cognitive systems, either.

Of course, if there were no way to identify some set of mechanisms as proprietary to the visual system, then one might be justified in denying the traditional distinction between perception and cognition. But we see no reason for such skepticism. In fact, we think that holding fixed (or abstracting away from) top-down effects provides one effective way of individuating perceptual systems. Having established a relatively plausible model of bottom-up visual processing one can thereafter look at how endogenous variables modulate that processing. Indeed, this appears to underlie the methodology employed by at least some cognitive neuroscientists.

Consider a study by Kok et al. (2013) in which participants implicitly learned two tone–orientation pairings. During subsequent trials, participants viewed random-dot-motion displays, where a subset of the dots moved in a coherent fashion. Participants were then asked to judge the direction of coherent motion. On trials when a tone was present, participants' orientation judgments showed a clear "attractive" bias – that is, they judged the orientation of the motion to be closer to the cued orientation than they did when there was no tone present (or a tone paired to a different orientation).

An important note: Participants performed this task within an fMRI scanner. The investigators then used a forward-modeling approach to estimate the perceived direction of coherent motion on each trial. This essentially involved collecting fMRI data from motion-selective voxels in areas V1, V2, and V3 on each trial, and using the data from the unbiased (bottom-up) trials to create an orientation-sensitive artificial neural network. The fMRI models for the biased (top-down influenced) trials turned out to match the participants' reports of the perceived direction better than they did the actual directions of motion, which suggests that the model accurately represents direction-sensitive processing in early vision. Further support for the validity of the model comes from the fact that there was a positive correlation between participants' behavioral and modeled responses. For example, if someone showed a stronger bias than others in the behavioral condition, then so did her fMRI forward model.

The moral we want to draw from this case is as follows. Artificial neural networks have long been used to model orientation processing in a *bottom-up fashion.* The network consists of neurons preferentially tuned to specific orientations, which in turn will exhibit differential activation patterns in the presence of coherent motion at different particular orientations. Kok et al. (2013) assume that such a model will predict people's behavioral responses in the absence of a tone *because* the system (comprising at least V1, V2, and V3) is specialized for processing visual inputs, and it will do so relying on bottom-up information alone in the absence of top-down modulation (which is what they found). In the presence of a tone, however, the model continues to match the behavioral response, but it no longer tracks the stimulus orientation. Thus, they infer that there must be some sort of top-down signal that alters the manner in which information is processed within the visual system.

In short, rather than obviating any distinction between perceptual and cognitive systems, this model seems to presuppose such a distinction, all the while claiming that vision is porous to endogenous information about the statistical regularities in the environment. In fact, it is in virtue of stable bottom-up models that one can begin to understand how top-down effects modulate visual processing. It may yet turn out that there *are no* (interesting) top-down modulations of the visual system, of course. That is an empirical possibility. But if there *are* interesting top-down effects (as in fact we think there are; see Ogilvie & Carruthers 2016), we don't think that should be regarded as especially revolutionary.

Studies on cognitively driven attention suggest that late vision is cognitively penetrated, whereas early vision is not

doi:10.1017/S0140525X15002484, e256

Athanassios Raftopoulos
Department of Psychology, University of Cyprus, 1678 Nicosia, Cyprus.
raftop@ucy.ac.cy

Abstract: Firestone & Scholl (F&S) examine, among other possible cognitive influences on perception, the effects of peripheral attention and conclude that these effects do not entail cognition directly affecting perception. Studies in neuroscience with other forms of attention, however, suggest that a stage of vision, namely late vision, is cognitively penetrated mainly through the effects of cognitively driven spatial and object-centered attention.

Firestone & Scholl (F&S) argue that the evidence that allegedly shows that cognition affects visual perception, once properly examined, does not support the view that the percept, the outcome of vision, is directly affected in a top-down way by cognition. By "direct cognitive effects," I mean, extending the authors' view (sect 4.5.1, para. 1), those cognitive influences that affect perceptual processing itself and change the percept, and not the cognitive effects that determine to what or where the perceiver attends. The authors state (sect. 1.2, para. 2) that whereas many previous discussions defended the modular nature of only a circumscribed, possibly unconscious, stage of visual processing – that is, "early vision" – they aim to assess the evidence for top-down effects on perception as a whole, including the conscious percept. Therefore, their discussion encompasses the cognitive effects on both early and late vision, the latter being the stage in which the percept is constructed.

Considering perception as a whole, the authors do not distinguish between cognitive effects on early vision and cognitive effects on late vision. This poses a problem because cognition affects early vision and late vision differently. I agree that early vision is cognitively impenetrable because there are no direct cognitive effects on early vision. There is, however, substantial neuropsychological evidence that late vision and, hence, the percept is directly cognitively penetrated because the perceptual processes of late vision use cognitive information as a resource and, thus, cognition modulates the processes of late vision.

The authors do not appreciate that late vision is cognitively penetrated because they restrict themselves to considering, among the various forms of attention, only peripheral attention – that is, attention as a determinant of the locus or object/feature of focus. The authors acknowledge that attentional effects that are not peripheral exist, and that attention may "interact in rich and nuanced ways with unconscious visual representations to effectively mold and choose a 'winning' percept – changing the content of perception rather than merely influencing what we focus on" (sect. 4.5, para. 6), but they opt to focus on peripheral attention.

Peripheral attention selects the input to perception but does not affect how the processing operates (ibid, sect. 4.5.1, para. 1). Accordingly, the authors correctly conclude that when perceptual behavior is explained by invoking the role of peripheral attention, even if peripheral attention is guided by cognitive states, that does not entail perception being cognitively penetrated, a view also shared by the majority of philosophers (Zeimbekis & Raftopoulos 2015). Attention, however, especially if viewed in line with the bias competition model, does not act only in this external way. Rather than merely selecting input, attention is integrated with, and distributed across, visual processing (Mole 2015; Raftopoulos 2009). Attentional effects are intrinsic in late vision rendering it cognitively penetrated (Raftopoulos 2009; 2011).

There are several ways cognition affects late vision, such as the application of concepts on some output of early vision so that hypotheses concerning the identities of distal objects can be formed and tested in order for the objects to be categorized and identified. Cognitively driven attention is one way – an often-studied way – cognition affects perceptions. Research suggests that cognitively driven spatial and object/feature-centered attention, as expressed by the N2 Event Related Potential (ERP) component, affects perceptual processing. The N2 is elicited about 200–300 ms after stimulus onset in monkeys and humans, in the area V4 and in the inferotemporal cortex. Research (Chelazzi et al. 1993; Luck 1995) suggests that the N2 reflects the allocation of attention to a location or object and is influenced by the type of the target and the density of the distractors. It is also sensitive to stimulus classification and evaluation (Mangun & Hilyard 1995). Thus, N2 is considered to be a component of cognitively driven or sustained attention.

Cognitively driven attention, for example, affects color processing in the human collateral sulcus (the main cortical area for analysis and coding of color information) at about 160 ms (Anllo-Vento et al. 1998). At about 235 ms in primate V1, attention distinguishes target from distractor curves (Roelfsema et al. 1998). Attention is thought to enhance the activity of neurons in the cortical regions that encode the attended stimuli. The timing of these cognitive effects places them within late vision but outside early vision, which means that late vision, but not early vision, is affected directly by cognition. It should be noted that the various precueing effects that affect early vision processing are not direct but indirect cognitive influences because they do not affect perceptual processing itself but, rather, the preparatory neuronal activity of the perceptual circuits (Raftopoulos 2015).

Why does attention enhance the activity of some neurons in the visual cortical regions during late vision? Clark (2013) argues that to perceive the world is to use what you know to explain away the sensory signal across multiple spatial and temporal scales; the process of perception is inseparable from cognitive processes. The aim of this interplay is to enable perceivers to respond and eventually adapt their responses as they interact with the environment so that this interaction be successful. Success in such an endeavor relies on inferring correctly (or nearly so) the nature of the source of the incoming signal from the signal itself.

Current research sheds light on the role of top-down cognitive effects in inferring correctly the identities of the distal objects during late vision. The cognitively driven direct attentional effects within late vision contribute to testing hypotheses concerning the putative distal causes of the sensory data encoded in the lower neuronal assemblies in the visual processing hierarchy. This testing assumes the form of matching predictions, made on the basis of a hypothesis, about the sensory information that the lower levels should encode assuming that the hypothesis is correct, with the current, actual sensory information encoded at the lower levels (Barr 2009; Kihara & Takeda 2010; Kosslyn 1994). To this aim, attention enhances the activity of neurons in the cortical regions that encode the stimuli that most likely contain information relevant to the testing of the hypothesis.

What draws the line between perception and cognition?

doi:10.1017/S0140525X15002617, e257

Martin Rolfs[a] and Michael Dambacher[b]

[a]*Bernstein Center for Computational Neuroscience & Department of Psychology, Humboldt University of Berlin, 10099 Berlin, Germany;*
[b]*Department of Psychology, University of Konstanz, 78457 Konstanz, Germany.*
martin.rolfs@hu-berlin.de
http://www.rolfslab.de/
michael.dambacher@uni-konstanz.de
https://scikon.uni-konstanz.de/personen/michael.dambacher/

Abstract: The investigation of top-down effects on perception requires a rigorous definition of what qualifies as perceptual to begin with. Whereas Firestone & Scholl's (F&S's) phenomenological demarcation of perception from cognition appeals to intuition, we argue that the dividing line is best attained at the functional level. We exemplify how this approach facilitates scrutinizing putative interactions between judging and perceiving.

In their target article, Firestone & Scholl (F&S) maintain the position that – for all we know – perception should be considered modular and impenetrable by cognitive top-down influences. They propose a recipe for the audit-proof identification of true top-down effects on perceptual function by avoiding six common pitfalls that have consistently undermined the significance of existing evidence. Given what is at stake – our understanding of the mind's fundamental architecture – we fully agree that this field requires the most rigorous empirical reasoning. However, to achieve this ambitious goal, we need to put our own house in order first and face the problem's other side: We need a clear definition of what qualifies as a perceptual phenomenon to begin with.

F&S appeal to the reader's intuition of what perception is and how it is different from cognition: "Just imagine looking at an apple in a supermarket and appreciating its redness (as opposed, say, to its price). *That* is perception" (sect. 1, para. 3). This phenomenological definition neatly illustrates to *us as observers* the quality of perception. Yet its amenability to *us as scientists* remains vague. Indeed, a purely phenomenological definition may expose perceptual measures to the pitfalls legitimately targeted by the authors. Asking an observer to appreciate the redness of an apple, for example, opens the judgment to coloring from memory or knowledge. Similarly, the conscious percepts that visual processing produces cannot be easily distinguished from the conscious cognitive state of the perceiver. Rather, the mere act of reading out the result of a perceptual process may stain its immaculacy, just as palpating a soft sponge will never reveal its true shape. The authors concur with this view, emphasizing the importance of performance-based measures tied directly to the perceptual phenomenon. To truly determine whether cognition penetrates perceptual processing, we need to know where perception ends and where cognition starts. How do we isolate perception empirically in the first place? How can we distinguish visual processing and experience from cognition to make its pure form – if it exists – amenable to empirical scrutiny?

The divide between perception and cognition is hard to maintain at the physiological level. F&S emphasize the importance of descending pathways on sensory areas of the brain, and indeed, a considerable number of physiological studies show effects of top-down knowledge at the earliest cortical stages of sensory processing (e.g., Boutonnet & Lupyan 2015; Dambacher et al. 2009; Kim & Lai 2012; Rabovsky et al. 2011; reviewed in Gilbert & Li 2013). In fact, anatomy tells us that the only substrates of visual processing that are not targeted by top-down feedback are in the retina, leaving little room to distinguish vision and cognition at this level of description.

We propose instead that perception is separated from cognition by its function. Perception has the purpose of providing packaged descriptions of the environment, which are then used by other functions of the mind, such as reasoning, conscious decision-making, or acting. To create these descriptions, perceptual processes extract stimulus features, group them in space and time, partition the scene into separate entities that obey figure–ground relationships, and label these objects or events. At this functional level, we argue, the distinction between perception and cognition works.

Agreeing on a level of description at which a dissociation between perception and cognition is justifiable is an important step. It allows us to identify the traces of processing in these functionally defined modules, which we can then use to track their cognitive malleability. But what would such traces be? To decide that, we contend, we need to turn to the properties of the perceptual system and identify those that uniquely serve its function and, thus, are unsuspicious to result from cognitive reasoning.

This idea is best illustrated using a tangible research example that entered the longstanding debate about whether the detection of causality in dynamic events results from perceptual processes (such as perceiving distance, motion, or color) or from cognitive reasoning that is based on the perceptual output. In a series of visual adaptation experiments, we showed that viewing many collision events in a rapid sequence (discs launching each other into motion) causes observers to judge subsequent events more often as noncausal (Rolfs et al. 2013). Critically, these negative aftereffects of exposure to causal events were retinotopic – that is, coded in the reference frame shared by the retina and early visual cortex – and were not explained by adaptation to other low-level features (e.g., motion, transient onsets, luminance, or contrast). A negative aftereffect (similar to those known in color vision or motion perception), its emergence from pure stimulus exposure, and – perhaps most important – its retinotopy are traces of visual functions. Arguably, their combination is an unlikely product of cognition. Therefore, these results strongly support the view that the detection of causal interactions is an achievement of the perceptual system, where visual routines in retinotopic brain areas detect and adapt to seemingly complex physical relations – cause and effect.

This example illustrates that the perceptual nature of a phenomenon becomes compelling when functional traces of the underlying sensory system can be revealed. Although we showcased the perceptual detection of causality, this is a general point that applies equally well to the study of established visual features such as motion or color. Evidence that perception is pliable by cognition becomes persuasive only if identifiable traces of perceptual processing follow observers' (allegedly biased) perceptual reports at every turn. We readily acknowledge that the feasibility of this approach depends on the particular research question and phenomenon, but we believe it complements the target article's call for eliminating confounding pitfalls. Yet, to facilitate decisive empirical contributions along this line, a rigorous definition of perception is in demand – one that is concrete enough to lend itself to the study of potential top-down effects of cognition. We suggested the functional level as expedient ground to evaluate the degree of isolation of perception from cognition. We challenge the authors to substantiate their definition of perception in this – or, if they disagree, in a different – realm.

Perception, cognition, and delusion

doi:10.1017/S0140525X15002691, e258

Robert M. Ross,[a,b] Ryan McKay,[a,b] Max Coltheart,[b] and Robyn Langdon[b]

[a]*Department of Psychology, Royal Holloway, University of London, Egham, Surry TW20 0EX, United Kingdom;* [b]*ARC Centre of Excellence in Cognition and its Disorders, and Department of Cognitive Science, Macquarie University, Sydney NSW 2109, Australia.*

robross45@yahoo.com.au
https://pure.royalholloway.ac.uk/portal/en/persons/robert-ross%286832d311-6807-48d4-8dba-c3ec998c55b9%29.html
ryanmckay@mac.com
https://pure.royalholloway.ac.uk/portal/en/persons/ryan-mckay_cda72457-6d2a-4ed6-91d7-cfd5904b91e4.html
max.coltheart@mq.edu.au
http://www.ccd.edu.au/people/profile.php?memberID=53
robyn.langdon@mq.edu.au
https://www.ccd.edu.au/people/profile.php?memberID=60

Abstract: Firestone & Scholl's (F&S) critique of putative empirical evidence for the cognitive penetrability of perception focuses on studies of neurologically normal populations. We suggest that a comprehensive exploration of the cognition–perception relationship also incorporate work on abnormal perception and cognition. We highlight the prominence of these issues in contemporary debates about the formation and maintenance of delusions.

> The matter of belief is, in all cases, different in kind from the matter of sensation or presentation, and error is in no way analogous to hallucination. A hallucination is a fact, not an error; what is erroneous is a judgment based upon it.
>
> — Russell (1914, p. 173)
>
> Perceiving is believing.
>
> — Fletcher & Frith (2009, p. 48)

Firestone & Scholl (F&S) present a stimulating critique of putative empirical evidence for the cognitive penetrability of perception. In making their case, however, they focus exclusively on research on perception and cognition in neurologically normal populations. In doing so, they neglect potentially important sources of informative data afforded by research on abnormal perception and cognition. Cognitive neuropsychology and cognitive neuropsychiatry are scientific disciplines that draw inferences about aspects of normal cognition (such as reading, object recognition, belief formation, reasoning, decision-making, and theory of mind) by studying patients with cognitive deficits (Coltheart 2007). We suggest that a comprehensive exploration of the relationship between perception and cognition should consider research from these disciplines. In particular, we demonstrate that the issue of cognitive penetrability looms large in contemporary debates about the formation and maintenance of delusions.

According to the two-factor theory of delusions, two distinct factors are causally responsible for the formation and maintenance of delusions (Coltheart et al. 2011). The first factor explains why the content of a delusional belief comes to mind, and the second factor explains why the belief is adopted rather than rejected. To date, the two-factor theory has focused on explaining specific monothematic delusions (delusions with one theme) associated with neurological damage, but some tentative suggestions have been made concerning how the two-factor theory might explain polythematic delusions (delusions with multiple themes) associated with psychiatric illnesses such as schizophrenia (Coltheart 2013).

Consider the Capgras delusion, a monothematic delusion in which a patient believes that a spouse or close relative has been replaced by an impostor. This delusion is thought to stem from disruption to the autonomic component of face recognition, such that familiar faces are recognized as familiar but *feel* unfamiliar. Empirical support for this hypothesis comes from studies that have found that, unlike control participants, patients with Capgras delusion do not show a pattern of autonomic discrimination (indexed by skin conductance response) between familiar and unfamiliar faces (e.g., Brighetti et al. 2007; Ellis et al. 1997; Hirstein & Ramachandran 1997). Other work, however, suggests that an anomalous autonomic response to familiar faces is not sufficient for the development of Capgras delusion. Tranel et al. (1995) studied patients with damage to ventromedial frontal regions of the brain who also failed to show a pattern of autonomic discrimination between familiar and unfamiliar faces, yet were not deluded.

According to two-factor theorists, Capgras patients and Tranel et al.'s (1995) ventromedial frontal patients share a common first factor: anomalous autonomic responses to familiar faces. What distinguishes them from one another is that Capgras patients have a second anomaly: a cognitive deficit in the ability to evaluate candidate beliefs. Analogous two-factor accounts have been offered for several other monothematic delusions (Coltheart et al. 2011). Importantly, all two-factor accounts are predicated on a conceptual distinction (and empirical dissociation) between perception and cognition: abnormal perception as the first factor and a cognitive belief evaluation deficit as the second factor. Furthermore, two-factor accounts are not committed to perception being cognitively penetrable, meaning that the two-factor theory is consistent with the hypothesis F&S present.

In contrast to the two-factor theory, the prediction-error theory of delusions holds that delusion formation and maintenance are caused by a single factor: aberrant processing of prediction errors (mismatches between expectations and actual inputs). In particular, delusions are conceived as attempts to accommodate inappropriately generated prediction error signals (Corlett et al. 2010; Fletcher & Frith 2009). Prediction-error theorists have tended to focus on delusions associated with schizophrenia, but they have also offered accounts of monothematic delusions associated with neurological damage (Corlett et al. 2010). Whereas the distinction between perception and cognition is critical for the two-factor theory, prediction-error theorists minimize or disavow this distinction:

> The boundaries between perception and belief at the physiological level are not so distinct. An important principle that has emerged is that both perception of the world and learning about the world (and therefore beliefs) are dependent on predictions and the extent to which they are fulfilled. This suggests that a single deficit could explain abnormal perceptions and beliefs. (Fletcher & Frith 2009, p. 51)

> Within this framework there is no qualitative distinction between perception and belief, because both involve making inferences about the state of the world on the basis of evidence. (Frith & Friston 2013, p. 5)

Furthermore, according to prediction-error theorists delusions provide examples of cognition penetrating perception: There exist "interactions between perception and belief-based expectation" (Corlett et al. 2010, p. 357), and "delusional beliefs can alter percepts such that they conform to the delusion" (Corlett et al. 2010, p. 353). This position seems to be in tension with the hypothesis F&S present. Consequently, we suggest it would be useful for F&S to expand the scope of their review by critically examining whether there is empirical evidence from research on delusions that cognition penetrates perception. If empirical evidence is compelling, then there exists a counterexample to F&S's hypothesis.

In this commentary, we have shown that the relationship between cognition and perception is a major point of interest in contemporary research on delusions. This suggests that evidence from cognitive neuropsychology and cognitive neuropsychiatry may play an important role in testing the hypothesis F&S present.

Attention and memory-driven effects in action studies

doi:10.1017/S0140525X15002575, e259

Philip Tseng,[a,b] Timothy Lane,[a,b,c,d] and Bruce Bridgeman[e]
[a]*Institute of Humanities in Medicine, Taipei Medical University, Taipei City 11031, Taiwan;* [b]*Brain & Consciousness Research Center, Shuang-Ho Hospital, Taipei Medical University, New Taipei City 23561, Taiwan;* [c]*Institute of European and American Studies, Academia Sinica, Taipei City 11529, Taiwan;* [d]*Research Center for Mind, Brain, and Learning, National Chengchi University, Taipei City 11605, Taiwan;* [e]*Department of Psychology, University of California, Santa Cruz, Santa Cruz, CA 95064.*
philip@tmu.edu.tw
http://medhuman1.tmu.edu.tw/philip/
timlane@tmu.edu.tw
http://chss.tmu.edu.tw/intro/super_pages.php?ID=intro3
bruceb@ucsc.edu
http://people.ucsc.edu/~bruceb/

Abstract: We provide empirical examples to conceptually clarify some items on Firestone & Scholl's (F&S's) checklist, and to explain perceptual effects from an attentional and memory perspective. We also note that action and embodied cognition studies seem to be most susceptible to misattributing attentional and memory effects as perceptual, and identify four characteristics unique to action studies and possibly responsible for misattributions.

Firestone & Scholl (F&S) make a strong case against the effect of top-down beliefs on perception. The argument for the cognitive impenetrability and modular nature of (visual) perception is reminiscent of the historic debate between Fodor and Turvey (Fodor & Pylyshyn 1981; Turvey et al. 1981), especially when most of the top-down modulation literature can find its roots in Gibson's (1966; 1979) ecological psychology. However, several aspects make F&S's article theoretically unique and important: (1) taking a logical approach such as with the El Greco fallacy, (2) speaking to a wider range of researchers beyond the action or embodied cognition literature, and (3) perhaps most important, providing a checklist of criteria for future studies against the six pitfalls they have cogently identified.

One common pitfall among studies that mistake top-down effects in judgment for perceptual effect is the use of subjective report in measuring percepts (e.g., light and darkness, reachability, distance). Not only is subjective report highly susceptible to task demand (e.g., Durgin et al. 2011a), but also it is problematic because it provides no additional information that would enable researchers to trace the source of the top-down effect. Accordingly, in order to dissociate perception and judgment, it is advisable to use performance-based measures that supply additional information (e.g., spatial, temporal), thereby making it possible to infer the stage of processing over which top-down cognition exerts its influence. One of our previous studies (Tseng & Bridgeman 2011) demonstrates this point: To test whether hands near a visual stimulus would *enhance* processing of the stimulus (as opposed to hands far; see Tseng et al. 2012 for a review), participants performed a forced-choice visual memory change detection task that provides accuracy and reaction time data, as opposed to subjective report. The rationale was that if hand proximity could really change the way visual stimuli are processed in a positive way, then hand proximity would predict enhanced visual processing, which would lead to better change detection performance. This would effectively rule out the judgment component; the participants cannot fake better performance.

Here it is important to clarify F&S's conceptual distinction between "perception and judgment": The two are not mutually exclusive, nor do they exhaust all alternatives (e.g., attention). Therefore, even if the judgment factor is accounted for by performance-based measures, such a result would not necessarily guarantee an effect in perception, especially because the effects on attention – which can modulate perception – can often disguise themselves as effects on perception (F&S's "periphery effect of attention"). To revisit the example above, although it is tempting to conclude that perception was directly modulated by hand proximity, it is equally plausible that the effect stemmed from biased attention near the hands. Indeed, analyzing participants' hit rates region by region on the screen showed a shift of correct responses toward the right-hand side, suggesting that the effect was mediated by biased spatial attention, not visual perception. This conclusion not only reemphasizes the importance of having a performance-based measure that can be analyzed differently to provide additional information, but also it is consistent with Pitfall 5, "peripheral attentional effects" (sect. 4.5), on F&S's checklist. The same rationale is also true for Pitfall 6, "memory and recognition" (sect. 4.6), and we attacked this problem by turning a potential artifact into an independent variable. Throwing a marble into a hole makes the thrower judge the hole as bigger following success than failure, but only if the hole is obscured after throwing. If the hole remains visible, the effect disappears (Blaesi & Bridgeman 2015; Cooper et al. 2012). The logic of this experiment is analogous to many efforts to demonstrate effects of action on perception, and it shows those results to affect memory, not perception. Modifying memory on the basis of experience is useful; modifying perception is not. Taken together, we recommend that future studies should consider Pitfalls 2, 5, and 6 together by controlling for judgment and memory effects and then moving on to tease apart the effects in perception versus attention.

Lastly, it is intriguing to us that a majority of the studies reporting top-down effects on perception are related to action (e.g., affordance, reachability). Might action studies be more susceptible to misattributing attentional or memory effects to perception? We speculate four possible reasons unique to the action literature for why this may be the case:

1. Universality: Due to motor action's depth in evolutionary time, action's effects on perception or attention are likely very widespread. This differs from the way in which ruminating about things, such as a sordid past, would make the room seem

darker (e.g., Banerjee et al. 2012; Meier et al. 2007). Because ruminations about the past involve parts of the cognitive economy that are evolutionarily recent, and because darkness metaphors of this type depend largely upon cultural interpretations that might be unique to humans, effects on perception are not as likely as actions and affordances.

2. Implicitness: Unlike certain, consciously accessible, top-down beliefs, information regarding action possibilities, or affordances, is often implicit properties that subjects may not be consciously aware of. Thus, the implicit nature of affordance information is assumed to be processed below consciousness threshold, and likely at the perceptual stage.

3. Well-established neurophysiology: The neuronal mechanisms for processing affordance or other action-relevant information (e.g., space, distance, graspability) have been well investigated in monkeys (e.g., Graziano & Botvinick 2002). Visual–tactile neurons in premotor and parietal cortices move their receptive fields with the hands instead of eyes, and they respond to objects that are within reach, even when "reachable" means "reachable with a tool."

4. Perception–action loop: The idea of perception–action coupling has been important in ecological psychology, and still is today in the embodied cognition literature. We suspect an overly literal interpretation of the idea can sometimes mislead researchers to mistake attentional effects as perceptual.

In summary, the effect of action on perception or attention is clearly quite different from other types of top-down beliefs. Although it is unfortunate that most action studies have mistaken attentional effects as perceptual, one can at least see why these studies may be more vulnerable to an inclination towards perceptual interpretations. Therefore, we recommend researchers in the field of perception and action and embodied cognition to especially consider F&S's arguments in the context of action when making conclusions.

ACKNOWLEDGMENT

This work is partially supported by a Taipei Medical University R&D Start-up Grant provided to Lane, and by Taiwan Ministry of Science & Technology Grants (104-2420-H-038-001-MY3, 105-2811-H-038-001, 105-2410-H-038-004, and 105-2632-H-038-001) also provided to Lane. It is as well supported by several grants to Tseng from Taiwan's Ministry of Science and Technology (104-2410-H-038-013-MY3), Taipei Medical University (TMU104-AE1-B07), and Shuang-Ho Hospital (105TMU-SHH-20).

Perception, as you make it

doi:10.1017/S0140525X15002678, e260

David W. Vinson,[a] Drew H. Abney,[a] Dima Amso,[b] Anthony Chemero,[c] James E. Cutting,[d] Rick Dale,[a] Jonathan B. Freeman,[e] Laurie B. Feldman,[f] Karl J. Friston,[g] Shaun Gallagher,[h] J. Scott Jordan,[i] Liad Mudrik,[j] Sasha Ondobaka,[g] Daniel C. Richardson,[g] Ladan Shams,[k] Maggie Shiffrar,[l] and Michael J. Spivey[a]

[a]*Cognitive and Information Sciences, University of California, Merced, Merced, CA 95340;* [b]*Department of Cognitive, Linguistic and Psychological Sciences, Brown University, Providence, RI 02912;* [c]*Department of Philosophy and Psychology, University of Cincinnati, Cincinnati, OH 45220;* [d]*Department of Psychology, Cornell University, Ithaca, NY 14850;* [e]*Department of Psychology, New York University, New York, NY 10003;* [f]*Psychology Department, University of Albany, SUNY, Albany, NY 12222;* [g]*Wellcome Trust Centre for Neuroimaging, University College London, London WC1E 6BT, United Kingdom;* [h]*Department of Philosophy, University of Memphis, Memphis, TN 38152;* [i]*Department of Psychology, Illinois State University, Normal, IL 61761;* [j]*School of Psychological Sciences, Tel Aviv University, Tel Aviv-Yafo, Israel;* [k]*Psychology Department, University of California, Los Angeles, Los Angeles, CA 90095;* [l]*Office of Research and Graduate Studies, California State University, Northridge, Northridge, CA 91330.*

dvinson@ucmerced.edu dabney@ucmerced.edu dima_amso@brown.edu chemeray@ucmail.uc.edu james.cutting@cornell.edu rdale@ucmerced.edu jon.freeman@nyu.edu lfeldman@albany.edu k.friston@ucl.ac.uk s.gallagher@memphis.edu jsjorda@ilstu.edu mudrikli@tau.ac.il s.ondobaka@ucl.ac.uk dcr@eyethink.org lshams@psych.ucla.edu mag@csun.edu spivey@ucmerced.edu

Abstract: The main question that Firestone & Scholl (F&S) pose is whether "what and how we see is functionally independent from what and how we think, know, desire, act, and so forth" (sect. 2, para. 1). We synthesize a collection of concerns from an interdisciplinary set of coauthors regarding F&S's assumptions and appeals to intuition, resulting in their treatment of visual perception as context-free.

No perceptual task takes place in a contextual vacuum. How do we know that an effect is one of perception *qua* perception that does not involve other cognitive contributions? Experimental instructions alone involve various cognitive factors that guide task performance (Roepstorff & Frith 2004). Even a request to detect simple stimulus features requires participants to understand the instructions (*language, memory*), keep track of them (*working memory*), become sensitive to them (*attention*), and pick up the necessary information to become appropriately sensitive (*perception*). These processes work in a dynamic parallelism that is required when one participates in any experiment. Any experiment with enough cognitive content to test top-down effects would seem to invoke all of these processes. From this task-level vantage point, the precise role of visual perception under strict *modular* assumptions seems, to us, difficult to intuit. We are, presumably, seeking theories that can also account for complex natural perceptual acts. Perception must somehow participate with cognition to help guide action in a labile world. Perception operating entirely independently, without any task-based constraints, flirts with hallucination. Additional theoretical and empirical matters elucidate even more difficulties with their thesis.

First, like Firestone & Scholl (F&S), Fodor (1983) famously used visual illusions to argue for the modularity of perceptual input systems. Cognition itself, Fodor suggested, was likely too complex to be modular. Ironically, F&S have turned Fodor's thesis on its head; they argue that perceptual input systems may interact as much as they like without violating modularity. But there are some counterexamples. In Jastrow's (1899) and Hill's (1915) ambiguous figures, one sees either a duck or rabbit on the one hand, and either a young woman or old woman on the other. Yet, you can cognitively control which of these you see. Admittedly, cognition cannot "penetrate" our perception to turn straight lines into curved ones in any arbitrary stimulus; and clearly we cannot see a young woman in Jastrow's duck-rabbit figure. Nonetheless, cognition can change our interpretation of either figure.

Perhaps more compelling are auditory demonstrations of certain impoverished speech signals called sine-wave speech (e.g., Darwin 1997; Remez et al. 2001). Most of these stimuli sound like strangely squeaking wheels until one is told that they are speech. But sometimes the listener must be told what the utterances are. Then, quite spectacularly, the phenomenology is one of listening to a particular utterance of speech. Unlike visual figures such as those from Jastrow and Hill, this is not a bistable phenomenon; once a person hears a sine wave signal as speech, he or she cannot fully go back and hear these signals as mere squeaks. Is this not top-down?

Such phenomena – the bistability of certain visual figures and the asymmetric stability of these speechlike sounds, among many others – are not the results of confirmatory research. They are indeed the "amazing demonstrations" that F&S cry out for.

Second, visual neuroscience shows numerous examples of feedback projections to visual cortex, and feedback influences on visual neural processing that F&S ignore. The primary visual cortex (V1) receives descending projections from a wide range of cortical areas. Although the strongest feedback signals come from nearby visual areas V3 and V4, V1 also receives feedback signals from V5/MT, parahippocampal regions, superior temporal parietal regions, auditory cortex (Clavagnier et al. 2004) and the amygdala (Amaral et al. 2003), establishing that the brain shows pervasive top-down connectivity. The next step is to determine what perceptual function descending projections serve. F&S cite a single paper to justify ignoring a massive literature accomplishing this (sect 2.2, para 2).

Neurons in V1 exhibit differential responses to the same visual input under a variety of contextual modulations (e.g., David et al. 2004; Hupé et al. 1998; Kapadia et al. 1995; Motter 1993). Numerous studies with adults have established that selective attention enhances processing of information at the attended location, and suppresses distraction (Gandhi et al. 1999; Kastner et al. 1999; Markant et al. 2015b; Slotnick et al. 2003). This excitation/suppression mechanism improves the quality of early vision, enhancing contrast sensitivity, acuity, d-prime, and visual processing of attended information (Anton-Erxleben & Carrasco 2013; Carrasco 2011; Lupyan & Spivey 2010; Zhang et al. 2011). This modulation of visual processing in turn supports improved encoding and recognition for attended information among adults (Rutman et al. 2010; Uncapher & Rugg 2009; Zanto & Gazzaley 2009) and infants (Markant & Amso 2013; 2016; Markant et al. 2015a). Recent data indicate that attentional biases can function at higher levels in the cognitive hierarchy (Chua & Gauthier 2015), indicating that attention can serve as a mechanism guiding vision based on category-level biases.

Results like these have spurred the visual neuroscience community to develop new theories to account for how feedback projections change the receptive field properties of neurons throughout visual cortex (Dayan et al. 1995; Friston 2010; Gregory 1980; Jordan 2013; Kastner & Ungerleider 2001; Kveraga et al. 2007b; Rao & Ballard 1999; Spratling 2010). It is not clear how F&S's theory of visual perception can claim that recognition of visual input takes place without top-down influences, when the activity of neurons in the primary visual cortex is routinely modulated by contextual feedback signals from downstream cortical subsystems. The role of downstream projections is still under investigation, but theories of visual perception and experience ought to participate in understanding them rather than ignoring them.

F&S are incorrect when they conclude that it is "eminently plausible that there are no top-down effects of cognition on perception" (final paragraph). Indeed, F&S's argument is heavily recycled from a previous BBS contribution (Pylyshyn 1999). Despite their attempt to distinguish their contribution from that one, it suffers from very similar weaknesses identified by past commentary (e.g., Bruce et al. 1999; Bullier 1999; Cavanagh 1999, among others). F&S are correct when they state early on that, "discovery of substantive top-down effects of cognition on perception would revolutionize our understanding of how the mind is organized" (abstract). Especially in the case of visual perception, that is exactly what has been happening in the field for these past few decades.

An action-specific effect on perception that avoids all pitfalls

doi:10.1017/S0140525X15002563, e261

Jessica K. Witt,[a] Mila Sugovic,[b] Nathan L. Tenhundfeld,[a] and Zachary R. King[a]

[a]*Department of Psychology, College of Natural Sciences, Colorado State University, Fort Collins, CO 80523;* [b]*Department of Psychological Sciences, College of Health and Human Sciences, Purdue University, West Lafayette, IN 47907.*

jessica.witt@colostate.edu
http://amplab.colostate.edu
milasugovic@gmail.edu **nate.tenhundfeld@colostate.edu**
zach.king@colostate.edu

Abstract: The visual system is influenced by action. Objects that are easier to reach or catch look closer and slower, respectively. Here, we describe evidence for one action-specific effect, and show that none of the six pitfalls can account for the results. Vision is not an isolate module, as shown by this top-down effect of action on perception.

> The plate. It looks so close. There are days when I first get out to the mound and it feels…like the plate is closer than it's supposed to be. Then I know right away. It's over. You are fucked. Fucked.
> — Pedro Martinez (Verducci 2000)

Hall-of-Fame baseball pitcher Pedro Martinez's experience can be explained by the action-specific account of perception. According to this account, people see the distance to or size of objects relative to their ability to act on these objects. At issue is whether supporting empirical findings reflect genuine effects on perception, or instead are a result of one of the six pitfalls Firestone & Scholl (F&S) outline. Fortunately, their claim that these issues have been "largely neglected" (sect. 4.4, para. 2) does not account for much empirical evidence directly addressing the issue with respect to action.

Their claim that *no* top-down effects on perception exist can be felled with the demonstration that one effect survives all pitfalls. We count four effects that meet this criterion. The first three are treadmill manipulations on perceived distance, reach-extending tools on perceived distance, and body-based manipulations in virtual reality on perceived size (see Philbeck & Witt 2015). We describe the fourth in detail.

In a paradigm known as Pong, participants attempted to catch a moving ball with a paddle that varied in size from trial to trial, and then estimated the speed of the ball. Previous research demonstrates that when participants play with a small paddle, the ball is harder to catch and is therefore subsequently judged to be moving faster than when they play with a big paddle (Witt & Sugovic 2010). Notably, paddle size influences perceptual judgments only when paddle size also impacts performance. When the ball is similarly easy to catch regardless of paddle size, the paddle has no effect on apparent speed (Witt & Sugovic 2012; Witt et al. 2012). These findings offer both disconfirmatory findings (Pitfall 1) and rule out low-level differences (Pitfall 4).

F&S criticized the term "perceptual judgments" as being vague and ambiguous. However, its use is frequently the researchers' acknowledgment that differentiating perception from judgment is nuanced and difficult. Indeed, F&S were unable to provide a scientific definition, instead relying too heavily on their own intuitions to distinguish perception and judgment (Pitfall 2). For example, comfort could very well be an affordance of an object that can be perceived directly (Gibson 1979). Nevertheless, the issue of distinguishing perception from judgment has been previously addressed. One strategy has been to use action-based measures for which no judgment is required. We modified the ball-catching task so that instead of continuously controlling the paddle, participants had only one opportunity per trial to move the paddle. Successful catches required precisely timing the action, and we analyzed this timing as an action-based measure of perceived speed. If the ball genuinely appears faster when the paddle is small, participants should act earlier than when the paddle is big. As predicted, participants acted earlier with the small paddle, indicating that the ball appeared faster, than with the big paddle (Witt & Sugovic 2013a). Because this measure is of action, and not an explicit judgment, the measure eliminates the concern of judgment-based effects (Pitfall 2). This measure also avoids the pitfall of relying on memory (Pitfall 6) because the action was performed while the ball was visibly moving.

Effects with action-based measures can also be taken as evidence against task demands (Pitfall 3). Additionally, we have directly measured participants' willingness to comply with task demands by purposefully inserting task demands into the design of the experiment. Participants were instructed on how to respond (e.g., to make sure to classify all fast speeds correctly), and we grouped participants based on their willingness to conform to these instructions. Importantly, both conforming and nonconforming participants showed identical action-specific effects of paddle size on apparent ball speed (Witt & Sugovic 2013b). The finding that nonconforming participants still show the same action-specific effect is evidence against a task demand explanation (Pitfall 3).

A final set of experiments explored the role of attention (Pitfall 5) in the Pong task by adding a secondary, attentionally demanding task (Witt et al. 2016). In one experiment, the secondary task was to count the number of flashes that occurred at the center of the screen. In another, the secondary task was to fixate on the ball and count the number of flashes that occurred on the ball as it moved across the screen. Regardless of attentional load location, paddle size continued to influence both perceptual judgments and action-based measures of ball speed. In other words, attention-based manipulations did nothing to diminish the action-specific effect; the effect of paddle size on apparent speed persisted in both cases. These studies rule out the final pitfall by showing that attention does not account for this particular action-specific effect.

We commend F&S for raising concrete concerns and future-oriented suggestions. We applied their checklist to one action-specific effect and found that none of the pitfalls could satisfactorily explain the effect of paddle size on apparent ball speed. We therefore conclude this effect is perceptual and demonstrates a genuine top-down influence on perception. Balls that are easier to catch are perceived to be moving slower than balls that are more difficult to catch. Going forward, researchers should apply this checklist to their own work to differentiate between effects that fall into the category of genuine perceptual effects and those that do not. However, the debate about whether there are any top-down effects on perception is decidedly in favor of a nonmodular view of vision.

Memory colours affect colour appearance

doi:10.1017/S0140525X15002587, e262

Christoph Witzel,[a,b] Maria Olkkonen,[c,d] and
Karl R. Gegenfurtner[a]

[a]*Abteilung Allgemeine Psychologie, Justus-Liebig-Universität Giessen, 35625 Giessen, Germany;* [b]*Laboratoire Psychologie de la Perception (LPP), Université Paris Descartes, 75006 Paris, France;* [c]*Department of Psychology, Durham University, Durham DH1 3LE, United Kingdom;* [d]*Institute of Behavioural Sciences, 00014 University of Helsinki, Finland.*

cwitzel@daad-alumni.de
http://lpp.psycho.univ-paris5.fr/person.php?name=ChristophW
maria.olkkonen@durham.ac.uk
https://www.dur.ac.uk/research/directory/staff/?mode=staff&id=14131
gegenfurtner@uni-giessen.de
http://www.allpsych.uni-giessen.de/karl

Abstract: Memory colour effects show that colour perception is affected by memory and prior knowledge and hence by cognition. None of Firestone & Scholl's (F&S's) potential pitfalls apply to our work on memory colours. We present a Bayesian model of colour appearance to illustrate that an interaction between perception and memory is plausible from the perspective of vision science.

When observers are asked to adjust an object with a typical colour (e.g., a yellow banana) to grey in an achromatic adjustment task, they adjust it slightly to the colour opposite to the typical colour (e.g., blue). This result implies that observers still perceive remnants of the typical colour of the object when the object is shown at a chromaticity that would be considered grey otherwise. And that shows that the knowledge about the typical colour of an object influences the perceived colour of that object (Hansen et al. 2006; Olkkonen et al. 2008; Witzel et al. 2011).

In contrast to earlier work on memory colour, including Duncker (1939) and Bruner et al. (1951), we particularly designed our achromatic adjustment method to circumvent problems related to judgement, memory, and response biases. It is important to note that Firestone & Scholl (F&S) did not correctly state our methods and findings. The banana was not "judged to be more than 20% yellow" (sect. 4.4.1, para. 3) at the neutral point; instead, observers needed to adjust the banana 20% in the "blue" direction to make it appear neutral. Yellow judgments would naturally be prone to judgement biases, whereas our nulling method is not, because participants are not asked to implicitly or explicitly rate the object colours. Instead, the achromatic adjustment task involves a genuinely perceptual comparison between the colour of the objects and the grey background to which the observers were adapted (Pitfall 2, "perception versus judgment," and Pitfall 6, "memory and recognition").

To avoid response biases, we presented the images in random colours at the beginning of each trial (Pitfall 3, "demand and response bias"). Doing so prevented a strategy of merely overshooting in the opposite colour direction, thus producing a spurious memory colour effect (Witzel & Hansen 2015). Even with this precaution, the observed effects went specifically in the opposite direction of the typical memory colours.

We carefully controlled our stimuli in their low-level, sensory characteristics (Pitfall 4, "low-level differences"). In contrast to F&S's general critique about the lack of control in luminance (sect. 4.4.1, para. 3), stimuli in the memory colour experiments were matched in average luminance (Hansen et al. 2006; Olkkonen et al. 2008; Witzel et al. 2011). Moreover, the control stimuli used to establish observer's grey adjustments independent of memory colour effects were matched in spatial and chromatic low-level properties with the colour-diagnostic images.

We also carefully explored the conditions under which the memory colour effect does not occur, providing "uniquely disconfirmatory predictions" (Pitfall 1, "an overly confirmatory research strategy," sect. 4.1). Objects without a memory colour and objects with achromatic (greyscale) memory colours, such as a striped sock and a white golf ball, do not produce any shift in grey adjustments (Witzel & Hansen, 2015; Witzel et al. 2011). Moreover, the effect lessens when decreasing characteristic features of the objects, such as in uniformly painted objects and outline shapes (Olkkonen et al. 2008; see also Fig. 1 in Witzel et al. 2011).

Finally, the task required observers to pay attention to the image in order to complete the grey adjustment, independent of whether the image showed a colour-diagnostic object or a control object (Pitfall 5, "peripheral attentional effects"). Apart from that, there is no reason a priori to assume that shifts of attention away from the stimulus would produce spurious memory colour effects.

We are left to explain why the greyscale image of the banana in the target article's Figure 2K does not appear yellow. The sensory signal coming from that figure unambiguously establishes that the colour difference between the leftmost and the rightmost banana is a difference between grey and yellow. The memory colour effect is more subtle and cannot compete with the unambiguous sensory information in Figure 2K (cf. our Fig. 1A). Contrary to Figure 2K, our method allows for detecting the small but systematic deviations of the grey perceived for example on a banana from the grey perceived on a control stimulus. These systematic deviations towards blue show that the recognition of the object as being a banana provides additional evidence for it being yellow that is combined with sensory evidence about the contrast between the adjusted colour and the grey background.

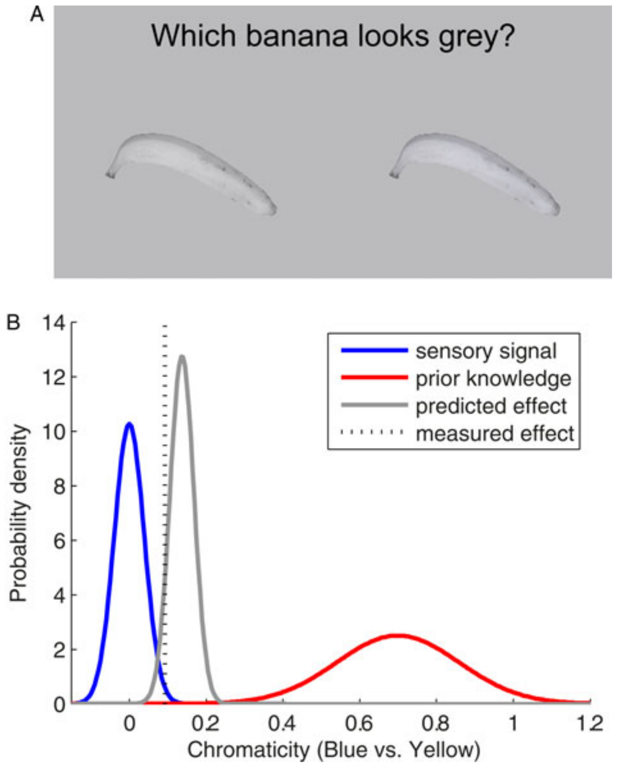


Figure 1 (Witzel et al.). (A) Illustration of the memory colour effect: the banana from Hansen et al. (2006) when it has the same chromaticity as the background (left) and when it has the average chromaticity that observers adjusted to make it appear grey (right). (B) Bayesian model of the memory colour effect. Hypothetical reliability of the sensory signal (blue line) and memory reliability (red line) for the typical yellow of a banana. The Bayesian combination of the two sources of information (grey line) predicts a shift in the perception of grey (at zero) towards yellow that corresponds to the memory colour effect. The observers compensate for this yellow shift in the percept (dotted vertical line) by adjusting the image towards blue.

In vision science, combining different types of evidence is most elegantly considered in a Bayesian framework (Maloney & Mamassian 2009). Consider our Figure 1b: When the images are achromatic, the sensory signal (blue curve) indicates greyness with a certain level of reliability. At the same time, prior knowledge about the typical colour of the object suggests that the object is likely to be coloured in its typical colour (red curve). Because sensory signals always contain uncertainty, combining sensory evidence with prior knowledge is a useful strategy to constrain perceptual estimates. As a result of the combination of sensory signals and prior knowledge in a Bayesian ideal observer model, the perceptual estimate of the colour (grey curve) shifts towards the typical colour of the object. When an observer is asked to make the object to appear grey, the colour setting needs to shift towards the opposite direction, thus producing the memory colour effect.

Whether memory colour effects are an example of top-down effects in the sense of *cognitive penetrability of perception* depends on the definition of perception and cognition (Witzel & Hansen 2015). We believe the notion that colour appearance is "low-level" whereas object recognition and memory are "high-level" (Eacott & Heywood 1995) is too simplified. In any case, evidence for the memory colour effects has also been observed in neuroimaging experiments (Bannert & Bartels 2013; Vandenbroucke et al. 2014) in early visual cortex, indicating that no matter at what stage they arise, they get propagated back to the early visual system.

The El Greco fallacy and pupillometry: Pupillary evidence for top-down effects on perception

doi:10.1017/S0140525X15002654, e263

Weizhen Xie and Weiwei Zhang
Department of Psychology, University of California, Riverside, Riverside, CA 92521.
weizhen.xie@email.ucr.edu **weiwei.zhang@ucr.edu**
http://sites.zanewzxie.org/
http://memory.ucr.edu

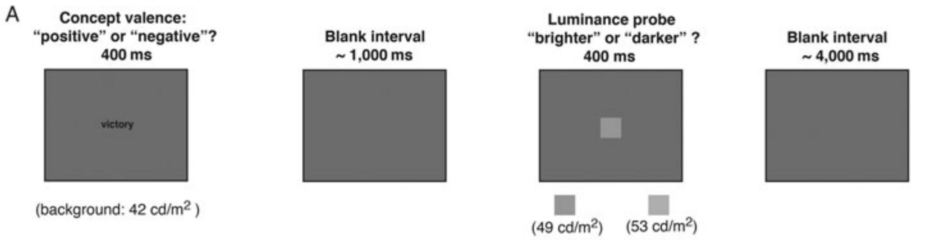
Abstract: In this commentary, we address the El Greco fallacy by reviewing some recent pupillary evidence supporting top-down modulation of perception. Furthermore, we give justification for including perceptual effects of attention in tests of cognitive penetrability. Together, these exhibits suggest that cognition *can* affect perception (i.e., they support cognitive penetrability).

Firestone & Scholl (F&S) argue against top-down influences of higher-level social cognitive factors (e.g., beliefs, desires, and emotion) on perception. They stipulate the conditions in which genuine top-down effects could be established and highlight a handful of pitfalls – some previous demonstrations of top-down effects in which the conditions were not satisfied.

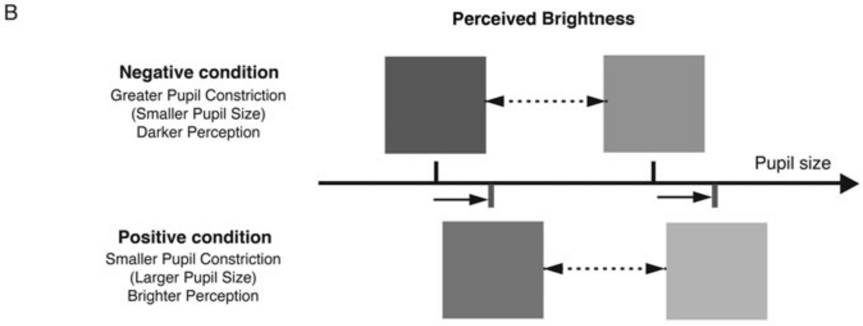
For example, F&S criticize a previous finding that positive (vs. negative) thoughts made the world look brighter (vs. darker, Meier et al. 2007). They rule out the possibility that these findings were results of cognitive penetration on perception. Specifically, in a task (Fig. 1A) modeled after Study 4 in Meier et al. (2007), participants discriminate between a darker and a brighter luminance probe following activation of emotional concepts using words with positive or negative meanings. If perception is modulated by emotional concepts, perceptual representations of the brighter probe and the darker probe should both shift rightward by positive concepts, resulting in indistinguishable luminance discriminability (i.e., *d′* in Signal Detection Theory, SDT) between the two luminance probes across emotion conditions (dashed lines in Fig. 1B). F&S thus argue that a genuine shift of perception by top-down factors could not manifest in behavioral reports (the El Greco fallacy, Pitfall 1). Therefore, any behavioral manifestation of changes in brightness perception induced by emotional concepts should result from response biases originating from postperceptual judgments (Firestone & Scholl 2014b) or low-level stimulus differences originating from bottom-up features (Firestone & Scholl 2015a; Lu et al. 2015).

While they have clearly demonstrated conceptual problems with the El Greco fallacy, F&S did not propose a solution for it. In the example of perceived brightness, a potential solution for this fallacy is to use direct or indirect assessment of perceived brightness, such as pupillometry, instead of relying on behavioral performance. Pupillary light response is traditionally believed to purely rely on bottom-up factors. However, some recent research has revealed robust cognitive effects on pupillary light responses (Hartmann & Fischer 2014; Laeng et al. 2012). That is, pupil size can be modulated by perceived brightness independent of physical brightness (e.g., Laeng & Endestad 2012; Laeng & Sulutvedt 2014; Mathôt et al. 2015; Naber & Nakayama 2013). For example, thinking about a bright event (e.g., a sunny day) leads to pupil constriction (Laeng & Sulutvedt 2014). These pupillary effects have been taken as evidence for cognitive penetrability (Hartmann & Fischer 2014), in that they are similar to pupillary responses to "real" visual perception induced by low-level physical stimuli.

Xie & Zhang (in preparation) generalized these pupillary effects in an experiment (Fig. 1A) modified from Study 4 in Meier et al. (2007). Accuracy in this experiment replicated the previous finding (Meier et al. 2007) that participants were more accurate in making a "brighter" response in the positive



The trial started with a brief presentation of a word at fixation, and participants reported whether the word had positive or negative meaning. After a blank interval for about a second, a gray square was presented at the center of the screen, and participants reported whether this luminance probe was a brighter one (53 cd/m^2) or a dimmer one (49 cd/m^2). The two probes were equally likely for each condition.



Positive thoughts evoked smaller pupil constriction for both luminance probes and consequently larger pupil size (arrows with solid lines) for brighter perception, than did negative thoughts, even though the perceptual discriminability between the two probes (dotted lines) were comparable across conditions.

Figure 1 (Xie & Zhang). Illustration of the task (A) and findings (B) from Xie & Zhang (in preparation).

condition than in a negative condition. However, perceptual discriminability of the two luminance probes (d', indicated by the dotted lines in Fig. 1B) was comparable between positive and negative conditions, as suggested by the El Greco fallacy, and participants were more inclined (more liberal response) to report brighter perception following positive thoughts than negative thoughts. These SDT measures alone seemed to suggest that effects of affective concepts on brightness perception were largely driven by response biases. However, these results were also consistent with a genuine shift in perceived brightness, as shown in Figure 1B, with a constant response criterion across the two conditions. Of these two possibilities, only the latter was supported by the pupil size data in that positive thoughts induced smaller pupil constriction for both luminance probes (arrows with solid lines in Fig. 1B), consequently resulting in larger pupil size for brighter perception (Chung & Pease 1999), than did negative thoughts. Note, the pupil effect here cannot be attributed to contextual priming or sensory adaptations. Together, these results supported, and more important provided a plausible mechanism for, the effects of emotional concepts on perceived brightness.

The pupil effect here refers to phasic changes in pupillary light response to the luminance probe, reflecting the transient sensory processing underlying the resulting perceived brightness. This phasic pupil size effect is different from tonic changes in pupil size elicited by affective concepts in Xie & Zhang (in preparation), in that tonic changes may result from processes that are not evoked by probe perception, such as arousal, task demand, and decisional uncertainty (Murphy et al. 2014). The tonic pupil size effects are therefore similar to differences in eye shapes (and thus pupil size) as intrinsic features of facial expressions in a previous study (Lee et al. 2014). F&S regard these tonic effects as changes in states of sensory organ (e.g., open vs. closed eye), and consequently F&S do not consider their effects on perception as evidence for cognitive penetrability. However, the phasic pupil size effects elicited by luminance probes are by no means changes in states of sensory organ, and therefore they warrant full consideration as candidates for evidence supporting cognitive penetrability.

Similar arguments can be made for perceptual effects of attention, which F&S simply attributed to changes in sensory inputs, instead of changes in sensory processing. This perspective seems to be an oversimplification. First, research on endogenous attention typically manipulates attention independent of eye movements (e.g., by presenting stimulus at fixation). The physical stimuli are thus kept constant between conditions, leading to the exact same optical inputs for sensory processing. Second, attention transiently modulates early feedforward sensory processing by amplifying sensory gain of attended information (Hillyard et al. 1998; Zhang & Luck 2009). It is therefore shortsighted to disregard perceptual effects of attention for cognitive penetrability.

In this commentary, we briefly reviewed some recent pupillary evidence supporting top-down modulation of perception and the justification for including attentional effects in tests of cognitive penetrability. Together, these pieces of evidence suggest that cognition *can* affect perception.

Authors' Response

Seeing and thinking: Foundational issues and empirical horizons

doi:10.1017/S0140525X16000029, e264

Chaz Firestone and Brian J. Scholl
Department of Psychology, Yale University, New Haven, CT 06520-8205.
chaz.firestone@yale.edu
brian.scholl@yale.edu

Abstract: The spectacularly varied responses to our target article raised big-picture questions about the nature of seeing and thinking, nitty-gritty experimental design details, and everything in between. We grapple with these issues, including the ready falsifiability of our view, neuroscientific theories that allow everything but demand nothing, cases where seeing and thinking conflict, mental imagery, the free press, an El Greco fallacy fallacy, hallucinogenic drugs, blue bananas, subatomic particles, Boeing 787s, and the racial identities of geometric shapes.

R1. Introduction

We are clearly not the only ones with strong views about how seeing relates to thinking. We were driven to explore the influences of cognition on perception primarily because this issue is so foundational to so many areas of cognitive science, and the commentaries on our target article exemplified this breadth and importance in several ways – hailing from many different fields (from systems neuroscience, to social psychology, to philosophy), drawing on vastly different methods (from rodent electrophysiology, to hue adjustment, to computational modeling), and originating from diverse perspectives (from predictive coding, to embodied cognition, to constructivism).

All of this led to a staggering diversity of reactions to our six "pitfalls," our conclusions about the state of the art, and our proposals for moving forward. Our approach was "theoretically unique" (**Tseng, Lane, & Bridgeman** [**Tseng et al.**]) but also "heavily recycled" (**Vinson, Abney, Amso, Chemero, Cutting, Dale, Freeman, Feldman, Friston, Gallagher, Jordan, Mudrik, Ondobaka, Richardson, Shams, Shiffrar, & Spivey** [**Vinson et al.**]); our recommendations constituted "an excellent checklist" (**Esenkaya & Proulx**) that was also "fundamentally flawed" (**Balcetis & Cole**); we gave a "wonderful exposé" (**Block**) that was also "not even wrong" (**Lupyan**); our critique was "timely" (**Gur**) but also "anachronistic" (**Clore & Proffitt**); we provided "a signal service to the cognitive psychology community" (**Cutler & Norris**) that was also "marginal, if not meaningless, for understanding situated behaviors" (**Cañal-Bruland, Rouwen, van der Kamp, & Gray** [**Cañal-Bruland et al.**]); we heard that "the anatomical and physiological properties of the visual cortex argue against cognitive penetration" (**Gur**), but also that our view "violates the functional architecture of the brain" (**Hackel, Larson, Bowen, Ehrlich, Mann, Middlewood, Roberts, Eyink, Fetterolf, Gonzalez, Garrido, Kim, O'Brien, O'Malley, Mesquita, & Barrett** [**Hackel et al.**]).

We are extremely grateful to have had so many of the leading lights of our field weigh in on these issues, and these 34 commentaries from 103 colleagues have given us a lot to discuss – so let's get to it. We first explore the foundational issues that were raised about the nature of seeing and its relation to thinking (sect. R2). Then, we take up the reactions to our article's empirical core: the six-pitfall "checklist" (sect. R3). Finally, we turn to the many new examples that our commentators suggested escape our pitfalls and demonstrate genuine top-down effects of cognition on perception (sect. R4).

R2. The big picture

Our target article was relentlessly focused on empirical claims, placing less emphasis on the broader theoretical landscape surrounding these issues. That focus was not an accident: We feel that purely theoretical discussions, though fascinating, have failed to move the debate forward. Nevertheless, many commentators raised issues of exactly this sort, and we have a lot to say about them.

R2.1. *See for yourself: Isolating perception from cognition*

Some commentators despaired over ever being able to separate seeing and thinking, denying that this distinction is real (**Beck & Clevenger**; **Clore & Proffitt**; **Goldstone, de Leeuw, & Landy** [**Goldstone et al.**]; **Hackel et al.**; **Keller**; **Lupyan**; **Miskovic, Kuntzelman, Chikazoe, & Anderson** [**Miskovic et al.**]; **Vinson et al.**) or well-defined (**Emberson**; **Gerbino & Fantoni**; **Rolfs & Dambacher**; **Witt, Sugovic, Tenhundfeld, & King** [**Witt et al.**]), and even claiming that "the distinction between perception and judgment, if there is one, is not clear and intuitive" in the first place (**Keller**).

R2.1.1. Seeing *versus* thinking. Speaking as people who can see and think (rather than as scientists who study perception and cognition), we find such perspectives baffling. One of the clearest and most powerful ways to appreciate that seeing and thinking *must* be different is simply to note that they often conflict: Sometimes, what you see is *different* than what you think. This conflict may occur not only because cognition fails to penetrate perception, but also because seeing is governed by different and seemingly idiosyncratic rules that we would never *think* to apply ourselves.

Perhaps nobody has elucidated the empirical foundations and theoretical consequences of this observation better than Gaetano Kanizsa, whose ingenious demonstrations of such conflict can, in a single figure, obliterate the worry that perception and cognition are merely "folk categories" (**Hackel et al.**) that "reify the administrative structure of psychology departments" (**Gantman & Van Bavel**) rather than carve the mind at its joints. For example, in Figure R1, you may *see* amodally completed figures that run counter to your higher-level intuitions of what *should* be behind the occluding surfaces, or that contradict your higher-level knowledge. Reflecting on such demonstrations, Kanizsa is clear and incisive:

> The visual system, in cases in which it is free to do so, does not always choose the solution that is most coherent with the context, as normal reasoning would require. This means that seeing follows a different logic – or, still better, that it does not perform any reasoning at all but simply works according to autonomous principles of organization which are not the same principles which regulate thinking. (Kanizsa 1985, p. 33)

Notice that Kanizsa's treatment forces us to acknowledge a *distinction* between seeing and thinking even before offering any *definition* of those processes. Indeed, literal definitions that cover all and only the relevant extensions of a concept are famously impossible to generate for anything worth thinking much about; by those austere standards, most *words* don't even have definitions. (Try it yourself with Wittgenstein's famous example of defining the word *game*.) So, too, for *perception*: It can't be that the scientific study of perception must be complete before we can say anything interesting about the

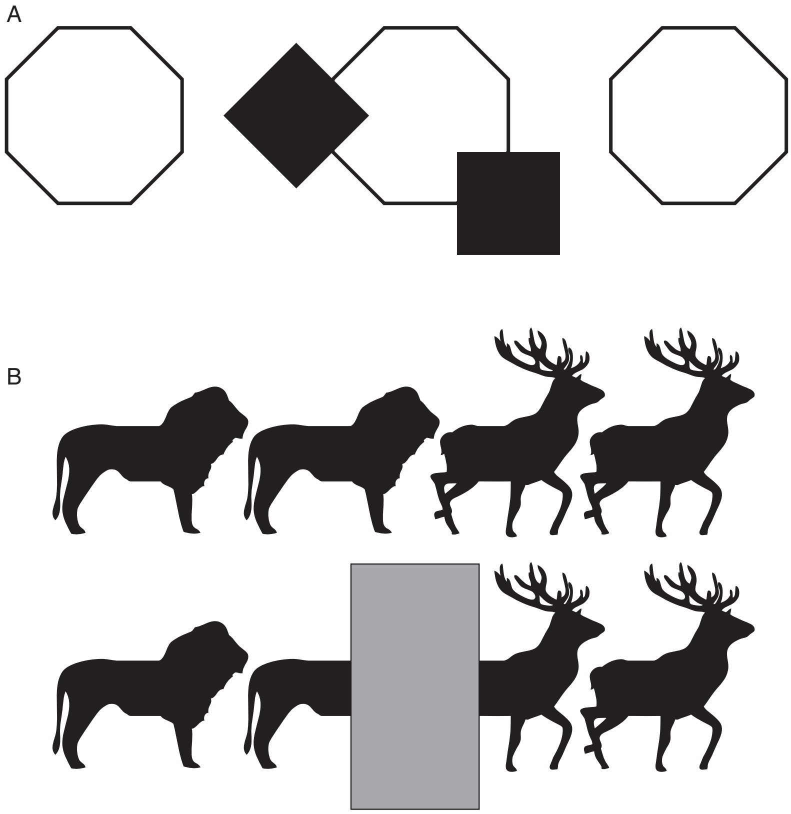


Figure R1 (Firestone & Scholl). Visual phenomena that contradict higher-level expectations. (A) Sandwiched by octagons, a partially occluded octagon looks to have an unfamiliar shape inconsistent with the scene (see also the discussion in Pylyshyn 1999). (B) An animal is seemingly stretched to an impossible length (and identity) when occluded by a wide surface. Adapted from Kanizsa and Gerbino (1982).

relationship between seeing and thinking. As Kanizsa's insights show, *distinctions* are what really matter. Such distinctions are thus precisely what our target article focused on, and we remain amazed that anyone looking at Figure R1 could deny the distinction between seeing and thinking.

Moreover, the concrete case studies we highlighted for our six pitfalls can serve a similar function as Figure R1. It can sometimes sound compelling in the abstract to question whether lines can be drawn between this or that process in the mind – for example, between perception and memory (e.g., **Emberson**; **Gantman & Van Bavel**; **Goldstone et al.**; **Lupyan**). But a concrete case study – for example, of "moral pop-out," which anchored Pitfall 6 ("Memory and Recognition") from our target article – wipes such abstract concerns away. And indeed, although the commentaries frequently discussed the distinction between perception and memory in the abstract – sometimes complaining that memory cannot be "cleanly split from perception proper" (**Lupyan**) – not a single commentator responded to this case study by rejecting the perception/memory distinction itself. (To the contrary, as we explore in sect. R3.6, Gantman & Van Bavel went to great lengths to argue that "moral pop-out" does not reflect semantic priming – presumably because they agreed that this alternative would indeed undermine their view.)

R2.1.2. Signatures of perception. Our target article focused primarily on case studies of phenomena that we argue *don't* reflect perception (instead involving processes such as higher-level judgment), but we have a robust theoretical and empirical interest in ways to show that various phenomena *do* reflect perception. Some commentators think our notion of perception is "extremely narrow" (**Cañal-Bruland et al.**) and "restrictive" (**Clore & Proffitt**), and that it "whittles the fascinating and broad domain of perception to sawdust" (**Gantman & Van Bavel**). We couldn't disagree more. Perception may be immune from cognitive influence, but it nevertheless traffics in a truly fascinating and impressively rich array of seemingly higher-level properties – including not just lower-level features such as color, motion, and orientation, but also causality (Scholl & Nakayama 2002), animacy (Gao et al. 2010), persistence (Scholl 2007), explanation (Firestone & Scholl 2014a), history (Chen & Scholl 2016), prediction (Turk-Browne et al. 2005), rationality (Gao & Scholl 2011), and even aesthetics (Chen & Scholl 2014). ("Sawdust"!)

Indeed, the study of such things is the primary occupation of our laboratory, and in general we think perception is far richer and *smarter* than it is often given credit for. But we don't think that anything goes, and we take seriously the need to carefully demonstrate that such factors are truly extracted during visual processing, per se. It is often difficult to do so, but it can be done – empirically and decisively (as opposed to only theoretically, as in proposals by **Halford & Hine** and **Ogilvie & Carruthers**). Indeed, our new favorite example of this was highlighted by **Rolfs & Dambacher**, who have demonstrated that the perception of physical causality (as when one billiard ball is seen to "launch" another) exhibits a property associated exclusively with visual processing: retinotopically specific adaptation (Rolfs et al. 2013; see also Kominsky & Scholl 2016) of the sort that also drives certain types of color afterimages. This example illustrates how "perception" can be identified – not by abstract definitional wordplay, but rather by concrete empirical signatures, of which there are many (for extensive discussion, see Scholl & Gao 2013).

R2.2. What would it take?

Several commentators worried that our view (as expressed in our target article's title) could not be disproven even in principle, and that it was even an "unfalsifiable tautology" (**Gerbino & Fantoni**). **De Haas, Schwarzkopf, & Rees** (**De Haas et al.**) challenged our view most directly in this way: "Specifically, what type of neural or behavioural evidence could refute it?" We accept this challenge.

We take our view to involve the most easily falsifiable claims in this domain in decades, and this assumption is absolutely central to our aims. **Gantman & Van Bavel** got things exactly right when they wrote that "the crux of F&S's argument lies in their empirical re-explanations of a handful of case studies. These are falsifiable." Similarly, we entirely agree with **Witt et al.** that our "claim that *no* top-down effects on perception exist can be felled with the demonstration that one effect survives all pitfalls."

So, what would it take to falsify our view in practice? That's easy: Every single one of the case studies discussed in our target article could easily have counted against our thesis! It could have been that when you give a good cover story for wearing a backpack (to mask its otherwise-obvious purpose), the backpack *still* makes hills look steeper. It could have been that when you blur faces, observers who don't see race also don't see the relevant lightness differences. It could have been that, under conditions characterized by an El Greco fallacy, the relevant top-down effects (e.g., of emotion on perceived brightness or of action-capabilities on perceived aperture width) disappear entirely, as they should. It could have been that when you carefully ask subjects to distinguish perception from "nonvisual factors," their responses clearly implicate perception rather than judgment. It could have been that "pop-out" effects in lexical decision tasks work for morality but *not* for arbitrary categories such as fashion. The list goes on.

Moreover, there's no "file-drawer problem" here; it's not that we've investigated dozens of alleged top-down effects and reported only those rare successes. Instead, every time we or others poke one of these studies with our pitfalls, it collapses. In other words, our view is eminently falsifiable, and indeed *we ourselves* – perhaps more so than any commentator – have tried our best to falsify it. We have simply failed to do so (and we engage with several new such claims in sect. R4).

R2.3. The perspective from neuroscience: Allowing everything but demanding nothing

Our discussion focused on *what we see* and *what we think*, and we suggested that perception is encapsulated from cognition. But many of the commentaries worried that in doing so we are living in the wrong century, harboring an "outdated view of the mind" (**Hackel et al.**). Instead, the more fashionable way to investigate what goes on in our heads is to consider "descending neural pathways" (**O'Callaghan, Kveraga, Shine, Adams, & Bar** [**O'Callaghan et al.**]), or "feedback projections" (**Vinson et al.**), or "reciprocal neural connections" (**Clore & Proffitt**), or a "dynamically reverberating loop" (**Miskovic et al.**), or a "continuum of brain modes" (**Hackel et al.**), or an "ongoing, dynamic network of feedforward and feedback activity" (**Beck & Clevenger**), or an "interconnected network of neurons, bathed in a chemical system, that can be parsed as a set of broadly distributed, dynamically changing, interacting systems" (**Hackel et al.**), or a "potpourri of synaptic crosstalk, baked into pluripotent cytocircuitry" (**OK, we made that one up**).

Many commentators noted, quite correctly, that we "readily dismiss the extensive presence of descending neural pathways" (**O'Callaghan et al.**) as having little to contribute to the core issue of how seeing and thinking interact. But we did so only in passing, in our zeal to focus on the relevant psychological experiments. And so in response, we will dismiss this work more comprehensively. We are, of course, aware of such ideas, but we think they are too often raised in these contexts in an uncritical way, and in fact are (some mixture of) irrelevant, false, and unscientific. Let's expand on this:

R2.3.1. "Unscientific." As far as we know, nobody thinks that every top-down effect of cognition on perception that *could* occur in fact *does* occur. For example, looking at Figure R2, you should experience the illusion of motion when you move your head from side to side (panel A) or forward and back (panel B), even though you can be morally certain that nothing is in fact moving in those images. (Indeed, to eliminate any doubt, you can view Figure R2 on a physical page, where a lifetime of experience and a vast body of knowledge about how ink and paper work can assure you that the images on the page are static.) Yet, as so often occurs with such phenomena, the illusion of motion persists. So, here is an example of what we know *failing* to influence what we see.

This sort of phenomenon invites a straightforward question for the perspectives articulated by so many of our neuro-inspired commentators: *How does this happen*? Given the overwhelming prevalence of loops and reentrance and descending pathways and interconnected networks and continua of brain modes, how is seeing insulated from thinking in this particular instance? Apparently, these rhapsodic accounts of the complete flexibility of perception are no obstacle to the thousands of visual phenomena that *aren't* affected by what we know, believe, remember, and so forth. As **Vinson et al.** concede, "Admittedly, cognition cannot 'penetrate' our perception to turn straight lines into curved ones."

In stark contrast with our view (which makes strong and necessary predictions; see sect. R2.2), these grand theories of brain function truly are unfalsifiable in the context of the present issues. Whenever there *is* an apparent top-down effect of cognition on perception, re-entrant/descending/recurrent … pathways/connections/projections get all of the credit. But whenever there *isn't* such an effect, nobody seems concerned, because that can apparently be accommodated just as easily. That is what an unfalsifiable theory looks like.

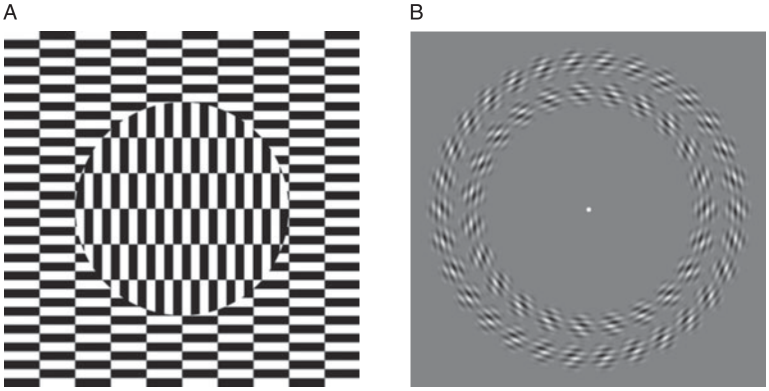


Figure R2 (Firestone & Scholl). Illusory motion in static images, defying our knowledge that these images are not, in fact, moving. (a) Moving one's eyes back and forth produces illusory motion in this illusion by Hajime Ouchi. (b) When the center is fixated, moving one's head toward and away from the image produces illusory rotary motion (Pinna & Brelstaff 2000; this version created by Pierre Bayerl and Heiko Neumann).

We have no doubt that these theories can contort themselves to explain away cases where thinking fails to affect seeing. (Maybe the solution has something to do with "the joint impact of priors and context-variable precision estimations" [**Clark**].) After all, many commentaries hedged their renditions of the pervasiveness of top-down processing in the brain, suggesting only that (with emphases added) "*much of* the neural hardware responsible for vision flexibly changes its function in complex ways depending on the goals of the observer" (**Beck & Clevenger**); that "the activity of neurons in the primary visual cortex is *routinely* modulated by contextual feedback signals" (**Vinson et al.**); and that "connectivity patterns within the cortex … *often* dominate" (**Hackel et al.**). But this is precisely the problem: Without independent constraints on "much of," "routinely," and "often," these brain models can accommodate any result. In short, they *allow everything*, but *demand nothing* – and so they don't explain anything at all. And until they are rid of this property, these "theories" are difficult to take seriously.

R2.3.2. "False." One commentator did take such views very seriously, issuing a powerful critique that met the ideas on their own terms. **Gur** makes a simple but ingenious case against various neuroscientifically inspired claims of cognitive penetrability, reaching the very opposite conclusion of so many others writing on this topic.

Perception, **Gur** notes, often traffics in fine-grained details. For example, we can perceive not only someone's face as a whole, but also the particular shape of a freckle on their nose. The only brain region that represents space with such fine resolution is V1 – which, accordingly, has cells with tiny receptive fields. So, to alter perception at the level of detailed perceptual experience, any influence from higher brain regions must be able match the same fine grain of V1 neurons. However, the only such connections – the "re-entrant pathways" trumpeted in so many commentaries – have much coarser resolution, in part because they pass through intermediate regions whose cells have much larger receptive fields (at least >5°, or about the size of your palm when held at arm's length). Therefore, influences from such higher areas cannot selectively alter the experience of spatial details smaller than the receptive fields of those cells. In other words, **Gur** concludes that the brain cannot even *implement* the top-down phenomena reported in the literature, making cognitive penetrability like "homeopathy … because no plausible mechanisms for its effects are suggested."

The key to such insights is critical: To appreciate the relevance (or lack thereof) of feedback connections in the brain, one must not only note their existence (as did so many commentaries), but also consider what they are *doing*. When you do only the former, you may hastily conclude that our view "violates the functional architecture of the brain" (**Hackel et al.**). But when you do the latter, you realize that "there is no feasible physical route for such a penetration" (**Gur**).

R2.3.3. "Irrelevant." Why is it so popular to leap from flexible models of brain function to a flexible relationship between seeing and thinking? **Desseilles & Phillips** may have shed some light on this issue: "Like the vast majority of professional neuroscientists worldwide, we consider that cognitions and perceptions are governed by specific patterns of electrical and chemical activity in the brain, and are thus essentially physiological phenomena."

But this line of reasoning is and has always been confused (see Carandini 2012; Fodor 1974). After all, it is equally true that cognition and perception are "governed" by the movement and interaction of protons and electrons; does this entail, in any way that matters for cognitive science, that seeing and thinking are *essentially subatomic phenomena*? That we should study subatomic structures to understand how seeing and thinking work? Clearly not.

Similarly, some commentaries seemed to go out of their way to turn our view into an easily incinerated straw man, alleging that "F&S assume that the words *cognition* and *perception* refer to distinct types of mental processes … localized to spatially distinct sets of neurons in the brain" (**Hackel et al.**). However, we assumed no such thing. Just as Microsoft Word is clearly encapsulated from Tetris regardless of whether they are distinguishable at the level of microprocessor structure, so too can perception be clearly encapsulated from cognition regardless of whether they are distinguishable at the level of brain circuitry (or subatomic structure).

R2.4. But why?

The core of our approach has been empirical, not theoretical, and we have avoided purely abstract discussions of why perception *should* or *should not* be encapsulated. Still, it can be interesting (if historically less productive) to consider *why* perception might be encapsulated from the rest of the mind.

R2.4.1. Flexibility, stability, and the free press. Many commentaries seemed to take for granted that cognitively penetrable vision would be a *good thing* to have, for example suggesting that a thoroughly top-down architecture is "undoubtedly a highly adaptive mechanism" (**O'Callaghan et al.**). This is not a foregone conclusion, however, as **Durgin** emphasized. Whereas other commentaries suggested that our target article "neglects the fundamental question what perception is for" (**Cañal-Bruland et al.**) – and that *action* is the answer – Durgin noted how successful action benefits from perceptual *stability*, because "momentary destabilization of space perception by desire, fatigue, and so forth would tend to undermine the whole point of perception as a guide for action" (Durgin).

These ideas relate to what our colleague Alan Gilchrist has informally called the "free press" model of perception. In government, as in the mind, it may serve certain short-term interests to actively distort the information reaching the people (or cognitive systems) who rely on it. However, in both cases, it is ultimately preferable not to exert such top-down influence, and instead to support a "free press" that can report what is happening honestly, without concern for what would be expedient at one particular time or for one special interest.

One reason to support the free press is that *one doesn't know in advance* just how this information will be used, and so any such distortions may have unintended negative consequences. In the case of perception, we may view a hill with a momentary intention to climb it; but even with this intention, we may also have other purposes in mind, for example using the hill as a landmark for later navigation, or to escape a flood. If the hill's perceived slant or height constantly shifts according to our energy levels or the weight on our shoulders (Bhalla & Proffitt 1999), then its utility for those other purposes will be undermined. (For example, we may think the hill offers greater safety from a flood than it truly does.) Better for vision to report the facts as honestly as possible and let the other systems relying on it (e.g., action, navigation, social evaluation, decision-making) use that information as they see fit.

R2.4.2. Protecting seeing from thinking. Another advantage of encapsulated perception is the benefit of automation. As another colleague, Scott Grafton, informally notes, encapsulation may sometimes seem like an exotic or specialized view when considered in the context of the mind, but it is actually commonplace in many control systems, whether engineered by natural selection or by people – and for good reason:

> Impenetrability … is the rule, not the exception. A pilot in a 787 gets to control a lot of things in his plane. But not everything. Much of it is now done by local circuits that are layered or protected from the pilot. A lot of plane crashes in modern planes arise when the pilot is allowed into a control architecture that is normally separate (fighting with the autopilot). Modern software in your computer keeps you out of the assembly code. For a human, how long do you think you would stay alive if you were allowed conscious control of your brainstem nuclei involved in blood pressure control, blood pH or cerebral perfusion pressure? Sensing and perception mechanisms likely operate with protocols that are not accessible by cognition. This should be the norm." (Grafton, personal communication)

In other words, by being encapsulated from thinking, seeing is *protected* from thinking. Our wishes, emotions, actions, and concerns are barred from altering visual processing so that they don't *mess it up*.

R3. The six pitfalls

The core of our target article explored how six concrete and empirically testable pitfalls can account for the hundreds of alleged top-down effects of cognition on perception reported in at least the last two decades – and how these pitfalls *do* account for many such effects in practice. Some commentaries accepted our recommendations, agreeing that "researchers should apply this checklist to their own work" (**Witt et al.**) and that "only research reports that pass (or at least explicitly address) F&S's six criteria can henceforth become part of the serious theoretical conversation" (**Cutler & Norris**). Other commentaries argued that our recommendations were "fundamentally flawed" (**Balcetis & Cole**). Here we respond to these many reactions.

R3.1. Pitfall 1: Uniquely disconfirmatory predictions

Although many commentators suggested that their favorite top-down effect escapes the El Greco fallacy, it was encouraging that nearly every commentary that discussed this pitfall seemed to accept its underlying logic: When the "measuring equipment" should be affected in the same way as whatever it's measuring, the effects must cancel out. Indeed, **Xie & Zhang** found this logic compelling enough to "fix" an El Greco fallacy that afflicted a previously reported top-down effect (Meier et al. 2007), though they ultimately implicated pupillary changes as the mechanism of the effect.

R3.1.1. An El Greco fallacy fallacy? One commentary, however, contested our application of the El Greco fallacy to a particular top-down effect. Holding a wide pole across one's body reportedly makes doorway-like apertures look narrower, as measured not only by adjusting a measuring tape to match the aperture's perceived width (Stefanucci & Geuss 2009), but also by adjusting a *second aperture* (Firestone & Scholl 2014b) – even though, according to the underlying theory, this second aperture should also have looked narrower and so the effects should have canceled out. **Hackel et al.** objected: "[T]he first aperture is meant to be passed through, whereas the second is not …. To assume that the top-down influence on width estimates would be the same and therefore cancel out under these distinct conditions suggests a misunderstanding of top-down effects."

However, it is **Hackel et al.** who have misunderstood both this top-down effect and the methodology of our El Greco fallacy studies. We did not blindly "assume" that the rod would influence both apertures equally – we actively built this premise into our study's design, anticipating exactly this concern. The original aperture-width study that inspired our own (Stefanucci & Geuss 2009) required subjects to imagine walking through the aperture before estimating its width, so as to engage the appropriate motor simulations. So, we made sure to ask our subjects to imagine walking through *both* apertures, on every trial, to ensure that both apertures would be "scaled" to the subject's aperture-passing abilities (Firestone & Scholl 2014b; Study 2). In other words, Hackel et al. simply have the facts wrong when they write that "the first aperture is meant to be passed through, whereas the second is not"; in truth, *both* apertures were viewed with passage in mind, just as the El Greco logic requires.

The fact that this crucial methodological detail (which was explicitly stated in the experiment's procedures) escaped the notice of all 16 of this commentary's authors amplifies one of our core themes: The empirical literature on top-down effects has suffered from a shortage of attention to exactly these sorts of details. Moving forward, it will not be enough to simply report an effect of some higher-level state on some perceptual property and leave it at that, without care to rule out other, nonperceptual interpretations. If there is one unifying message running through our work on this topic, it is this: The details matter.

R3.2. Pitfall 2: Perception versus judgment

Our target article called for greater care in distinguishing perception from postperceptual judgment. For example, subjects who are asked how far, large, or fast some object is might respond not only on the basis of how the object *looks*, but also on the basis of how far or large or fast they *think* it is. Many commentators accepted this distinction and our suggestions for exploring it. However, multiple commentaries denied that this pitfall afflicts the research we discussed, because of special measures that allegedly rule out judgment conclusively.

R3.2.1. Does "action" bypass judgment? At least two commentaries (**Balcetis & Cole**; **Witt et al.**) argued that so-called "action-based measures" can rule out postperceptual judgment as an alternative explanation of alleged top-down effects. For example, rather than verbally reporting how far away or how fast an object is, subjects could throw a ball (Witt et al. 2004) or a beanbag (Balcetis & Dunning 2010) to the object, or catch the object if it is moving (Witt & Sugovic 2013b). Witt et al. asserted that such measures directly tap into perception and not judgment: "Because this measure is of action, and not an explicit judgment, the measure eliminates the concern of judgment-based effects."

But that assertion is transparently false: In those cases, actions not only can *reflect* explicit judgments, but also they often *are* explicit judgments. This may be easier to see in more familiar contexts where perception and judgment come apart. For example, objects in convex passenger-side mirrors are famously "closer than they appear," and experienced drivers learn to account for this. Someone looking at an object through a mirror that they know distorts distances may *see* an object as being, say, 20 feet away, and yet *judge* the object to be only 15 feet away. Indeed, such a person might respond "15 feet" if asked in a psychology experiment how far away they think the object is. What about their actions? According to **Witt et al.**'s line of argument, once people are asked to *throw a ball* at the object, they will somehow forget everything they know about the mirror's distortion and simply throw the ball as far as the object looks, without correcting for that higher-level knowledge. But that seems absurd: Our object-directed actions can and do incorporate what we think, know, and judge – in addition to what we see – and there is no reason to think that the actions in Witt et al.'s various experiments are any different.[1]

R3.3. Pitfall 3: Task demands and response bias

Certain points of disagreement with our commentators were not unexpected. For example, we anticipated having to defend the distinctions we drew between perception, attention, and memory (see sect. 3.5 and 3.6). We were genuinely surprised, however, that a few brave commentators rejected our recommendations about controlling for task demands (**Balcetis & Cole**; **Clore & Proffitt**). We suggested that many apparent top-down effects of cognition on perception arise because subjects figure out the purpose of such experiments and act compliantly (cf. Orne 1962), or otherwise respond strategically,[2] and so we made what we thought were some mild recommendations for being careful about such things (e.g., actively asking subjects about the study's purpose, and taking measures to mask the purpose of the manipulations).

Balcetis & Cole rejected these recommendations as "fundamentally flawed," and replaced them with "five superior techniques" of their own (see also **Clore & Proffitt**). Though we were happy to see these concrete details discussed so deeply, these new recommendations are no substitute for the nonnegotiable strategies of masking demand and carefully debriefing subjects about the experiment – and in many cases these supposedly "superior" techniques actually *worsen* the problem of demand.

R3.3.1. Asking.... A primary technique we advocated for exploring the role of demand in top-down effects is simply to *ask the subjects* what they thought was going on. This has been revealing in other contexts: For example, more than 75% of subjects who are handed an unexplained backpack and are then asked to estimate a hill's slant believe that the backpack is meant to alter their slant estimates (stating, e.g., "I would assume it was to see if I would overestimate the slope of the hill"; Durgin et al. 2012). It is hard to see what could be flawed about such a technique, and yet it is striking just how few studies (by our count, zero) bother to systematically debrief subjects as Durgin et al. (2009; 2012) show is necessary.

Clore & Proffitt suggested that asking subjects about their hypotheses once the experiment is over fails to separate hypotheses generated during the task from hypotheses generated post hoc. We are unmoved by that suggestion. First, the same studies that find most subjects figure out the experiment's purpose and change their estimates also find those same subjects are driving the effect (Durgin et al. 2009). But second, **Clore & Proffitt**'s observation makes debriefing a *stronger* test: If your effect is reliable even in subjects who don't *ever* guess the experiment's purpose – whether during or after the experiment – then that is even more compelling evidence against a role for demand. There is no reason not to ask.

R3.3.2. …and telling. **Balcetis & Cole** specifically criticized our recommendation to use cover stories to mask the purpose of manipulations (e.g., telling subjects that a backpack carried electrodes or that a pole was for balance): "Alternative cover stories do not remove the opportunity to guess hypotheses, nor do they eliminate the possibility that participants will amend their responses in accordance with their conjectured suppositions. They simply introduce new task demands."

We agree that there is no such thing as a demand-free environment, but what is the problem here? Alternative cover stories could be problematic only if the "conjectured suppositions" they imply would produce a *directional* bias in estimates. What is the implied direction in telling subjects that a backpack contains electrodes (Durgin et al. 2009) or that a pole is for keeping one's balance (Firestone & Scholl 2014b)? These cover stories *eliminated* the relevant effects; they didn't *reverse* them. Giving a cover story for the backpack, for example, led subjects to make the same slant estimates they made with no backpack at all. Is it really **Balcetis & Cole**'s contention that when backpack-wearing subjects were given a cover story, they saw the slope as steeper but then intentionally *lowered* their estimates (for some unarticulated reason), and by precisely the amount required to make it look like there was no effect at all? That is the *only* possibility that would undermine the use of alternative cover stories, and accordingly we find the surprising resistance to this invaluable methodological technique to be uncritical and unfounded.

R3.3.3. Flawed alternatives. In place of cover stories and careful debriefing, **Balcetis & Cole** suggested five alternative techniques to combat demand. We briefly respond to each:

1. *Accuracy incentives*, including paying subjects extra money for correct responses: This technique sounds promising in principle, but it has foundered in practice. Balcetis and Dunning (2010), for example, told subjects they could win a gift card by throwing a beanbag closer to it than any other subject; subjects made shorter throws to a $25 gift card than a $0 gift card, suggesting that desirable objects look closer. But subjects care about winning valuable gift cards (and not worthless ones), and so they may have been differently engaged across these situations, or used different strategies. Indeed, follow-up studies showed that such strategic differences alone produce similar effects (Durgin et al. 2011a).[3]

2. *Counterintuitive behavioral responses,* including standing farther from chocolate if it looks closer (Balcetis & Dunning 2010): Whether something is intuitive or counterintuitive is an empirical question, and one cannot be sure without investigation. Rather than potentially underestimating subjects' insights, we recommend *asking them what they thought*, precisely to learn just what is (counter)intuitive.

3. *Between-subjects designs*, so as not to highlight differences between conditions: such designs can help, but they are completely insufficient. The backpack/hill study, for example, employed a between-subjects design, and subjects readily figured out its purpose anyway (Bhalla & Proffitt 1999; Durgin et al. 2009; 2012).

4. *Double-blind hypothesis testing:* This is another good idea, but experimenter expectancy effects are different than task demands. Our concern is not that subjects may divine the study's purpose from the experimenter's behavior; it is that *the task itself* makes the purpose transparent. The simple act of giving subjects an unexplained backpack and asking them to estimate slant reveals the hypothesis no matter what the experimenter knows or doesn't know.

5. *Dissociate measures from demand*, for example by having subjects throw a beanbag to a $100 bill they can win in a later unrelated contest (Cole & Balcetis 2013): Again, this may be helpful in principle, but in practice it may cause more problems than it solves. Indeed, that same study also showed that subjects felt more excited or "energized" upon seeing the winnable $100 bill (compared with no bill) – and that sort of confounding factor could independently influence subjects' throws.

In short, we reject the contention that **Balcetis & Cole**'s alternatives are "superior" – or even remotely sufficient. If an empirical paper implemented only their techniques, we would be entirely unconvinced – and you should be, too. The direct approach is the truly superior one: Cover stories have proven effective in exactly these circumstances, and they can and should be used in an unbiased way. And ever since Durgin et al. (2009), asking subjects what they think is simply *mandatory* for any such experiment to be taken seriously.

R3.4. Pitfall 4: Low-level differences (and amazing demonstrations!)

Low-level differences in experimental stimuli (e.g., shading, curvature) can be confounded with higher-level differences (e.g., race, being an animal/artifact; Levin & Banaji 2006; Levin et al. 2001), such that it is not always clear which is responsible for an apparent top-down effect. We showed that low-level factors *must* contribute to one such effect: African-American faces look darker than Caucasian faces even when the images are equated for mean luminance (Levin & Banaji 2006); however, when the faces are blurred, even subjects who do not appreciate race in the images still judge the African-American face to be darker than the Caucasian face (Firestone & Scholl 2015a), implying that low-level properties (e.g., the *distribution* of luminance) contribute to the effect.

Levin, Baker, & Banaji (**Levin et al.**) engaged with this critique in exactly the spirit we had hoped, and we thank them for their insightful and constructive reaction. However, we contend that they have misinterpreted both our data *and theirs*.

R3.4.1. Seeing race? **Levin et al.**'s primary response was to suggest that subjects could still detect race after our blurring procedure, reporting above-chance performance in race identification in a two-alternative forced-choice (2AFC) between "Black" and "White." But this response simply misunderstands the logic of our critique, which is not that it is completely impossible to guess the races of the faces, but rather that even those subjects who *fail* to see race in the images still show the lightness distortion. Our experiment gave subjects every opportunity to identify race in the images: We asked them to (a) describe the images in a way that could help someone identify the person; (b) explicitly state whether the races of the faces looked the same or different; (c) explicitly categorize the faces from a list of possible races; and (d) tell us if they ever thought about race but had been embarrassed to say so. Even those subjects who repeatedly showed no evidence of seeing race in the images (and indeed, even those subjects who explicitly thought the two images were *of the same person*) still judged the blurry African-American face to be darker.

Worse yet, 2AFC tasks are notoriously unreliable for higher-level properties such as race, which is why our own studies did not use them. **Levin et al.** concluded from above-chance performance in two-alternative racial categorization that "the blurring left some race-specifying information in the images." But when you give subjects only two options, they can choose the "right" answer for the *wrong reason* or be prompted to look for information that hadn't previously seemed relevant – for example particular patterns of shading that they hadn't previously considered in a racial context.

For example, suppose that instead of blurring the images, we just replaced them with two homogeneous squares, one black and one white, and then we adapted **Levin et al.**'s paradigm to those images – so that the question was "Using your best guess, how would you differentiate these *squares* by race?" – and subjects had to choose which *square* was "African-American" or "Caucasian" (Fig. R3). In fact, we made this thought experiment an empirical reality, using 100 online subjects and the same parameters as Levin et al. *All 100 subjects* chose "African-American" for the black square and "Caucasian" for the white square.[4] Do these results imply that subjects perceived race in these geometric shapes? Does it mean that "replacing the faces with homogeneous squares left some race-specifying information in the images"? Obviously not – but this is the same logic as in Levin et al.'s commentary. The mere ability to assign race when forced to do so doesn't imply that subjects actively categorized the faces by race; our data are still the only investigation of this latter question, and they suggest that even subjects who don't categorize the faces as African-American and Caucasian still experience distorted lightness.

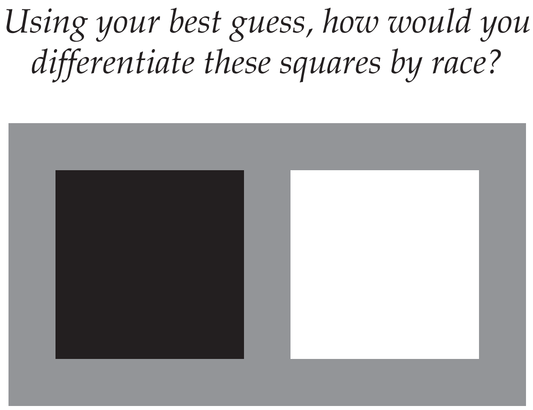


Figure R3 (Firestone & Scholl). Two-alternative forced-choice judgments can produce seemingly reliable patterns of results even when subjects don't base their judgments on the property of interest. If you *had to choose*, which of these squares would you label "African-American," and which would you label "Caucasian"?

R3.4.2. Other evidence. Levin et al. correctly note that we discussed only one of Levin and Banaji's (2006) many experiments in our target article, and they suggest that the other data (e.g., with line drawings equated for the distribution of luminance) provide better evidence. But we focused on Levin and Banaji's (2006) "demo" rather than their experiments not because it was easy to pick on, but rather because it was the most compelling evidence we had ever seen for a top-down effect – much more so than their other experiments, which suffered from an El Greco fallacy, weren't subjectively appreciable, and included a truly unfortunate task demand (in that subjects were told in advance that the study would be about "how people perceive the shading of faces of different races," which may have biased subjects' responses). By contrast, their "demo" seemed like the *best* evidence they had found, and so we focused on it. A major theme throughout our project has been to focus not on "low-hanging fruit" but instead on the strongest, most influential, and best-supported cases we know of for top-down effects of cognition on perception. We happily include Levin and Banaji's (2006) inspiring work in that class.

R3.5. Pitfall 5: Peripheral attentional effects

Most commentaries agreed that peripheral effects of attention (e.g., attending to one location or feature rather than another) don't "count" as top-down effects of cognition on perception, because – like shifts of the eyes or head – they merely select the *input* to otherwise-impenetrable visual processing. (Most, but not all: **Vinson et al.** suggested that our wide-ranging and empirically anchored target article is undermined by the century-old duck–rabbit illusion. Believe it or not, we knew about that one already – and it, like so many other ambiguous figures, is easily explained by appeal to attentional shifts; Long & Toppino 2004; Peterson & Gibson 1991; Toppino 2003).

Other commentaries, however (especially **Beck & Clevenger**; **Clark; Goldstone et al.**; **Most**; **Raftopoulos**), argued that attention "does not act only in this external way" (Raftopoulos). Clark, for example, pointed to rich models of attention as "a deep, pervasive, and entirely *non-peripheral* player in the construction of human experience," and asked whether attention can be written off as "peripheral."

We are sympathetic to this perspective in general. That said, we find allusions to the notion that attention can "alter the balance between top-down prediction and bottom-up sensory evidence at every stage and level of processing" (**Clark**) to be a bit too abstract for our taste, and we wish that these commentaries had pointed to particular experimental demonstrations that they think could be explained only in terms of top-down effects. Without such concrete cases, florid appeals to the richness of attention are reminiscent of the appeals to neuroscience in section R2.3: They sound compelling in the abstract, but they may collapse under scrutiny (as in sect. R2.3.2).

In general, however, our claim is not that *all* forms of attention must be "peripheral" in the relevant sense. Rather, our claim is that *at least some* are merely peripheral, and that many alleged top-down effects on perception can be explained by those peripheral forms of attention. This is why **Lupyan** is mistaken in arguing that "Attentional effects can be dismissed if and only if attention simply changes input to a putatively modular visual system"; attention may be a genuine alternative explanation just as long as attention *sometimes* changes input to later visual processing – because then such attentional effects must be actively ruled out by careful experimental tests of the sorts sketched in our target article.

R3.5.1. On which side of the "joint" is attention? So, what about those cases of attention that *aren't* like moving your eyes? To be sure, we think such cases are rarer than many commentaries imagine. For example, attending to features, rather than locations, may not be analogous to moving one's eyes, but it is importantly analogous to seeing through a tinted lens – merely increasing sensitivity to certain features rather than others. Across the core cases of attending to locations, features, and objects, both classical and contemporary theorizing understands that, fundamentally, "attention is a *selective* process" that modulates "early perceptual *filters*" (Carrasco 2011, pp. 1485–1486, emphasis added). That is what we mean when we speak of attention as constraining input: Attention acts as a *filter* that *selects* the information for downstream visual processing, which may itself be impervious to cognitive influence.

However, even if attention can go beyond this role and "alter the balance between top-down prediction and bottom-up sensory evidence at every stage and level of processing" (**Clark**), we find it odd to move from such sophisticated attentional processing to the further claim that perception is "cognitively penetrated" by attention (**Raftopoulos**). The controversy over top-down effects of cognition on perception is a controversy over the *revolutionary* possibility that what we see is directly altered by how we think, feel, act, speak, and so forth. But attention's role in perception simply cannot be revolutionary in this way: As **Block** noted in his commentary, "attention works via well-understood *perceptual mechanisms*" (emphasis his); and, as he has noted more informally, attention – unlike morality and hunger, say – is already extensively studied by vision science, and it fits comfortably within the orthodox framework of how the mind (in general) and perception (in particular) are organized. Our project concerns the "joint" between perception and cognition, and attention unquestionably belongs on the *perception* side of this joint. If some continue to think of attention as a nonperceptual influence on what we see, they can do so; but to quote Block out of context, "If this is cognitive penetration, why should we care about cognitive penetration?"

R3.6. Pitfall 6: Memory and recognition

Although many commentaries discussed the distinction between perception and memory – some suggesting that memory accounts for even *more* top-down effects than we suggested (**Tseng et al.**) – two commentaries in particular protested our empirical case studies of this distinction (perhaps unsurprisingly, given that those case studies involved their work; **Gantman & Van Bavel**; **Lupyan**). At the same time, these commentaries sent mixed signals: Both objected to our distinction between perception and memory, claiming that it "carves the mind at false joints" (**Gantman & Van Bavel**) because memory cannot be "cleanly split from perception proper" (**Lupyan**); but both then went to extraordinary lengths to try to *rule out* the memory-based interpretations we offered – apparently agreeing that such alternatives would undermine their claims. How compelling were these attempted rebuttals?

R3.6.1. "Moral pop-out" does not exist. Moral words are identified more accurately than random nonmoral words, which led Gantman and Van Bavel (2014) to claim that "moral concerns shape our basic awareness" (p. 29) – a claim that has since been upgraded to "human perception is preferentially attuned to moral content" (Gantman & Van Bavel 2015, p. 631; though, see Firestone & Scholl 2016). However, the moral words in these studies were semantically related to each other (e.g., *crime*, *punishment*), whereas the nonmoral words were not (e.g., *steel*, *ownership*), which led us to suspect that semantic priming due to spreading activation – a phenomenon of memory rather than perception – might explain the effect. Sure enough, you can obtain "pop-out" effects with *any* arbitrary category of related words (Firestone & Scholl 2015b), including fashion (e.g., *blouse*, *dress*; pop-out effect: 8.6%) and transportation (e.g., *car*, *bus*; pop-out effect: 4.3%). (We also replicated the effect with morality; pop-out effect: 3.9%, which matched Gantman & Van Bavel's original report.) Moreover, although our experiments were not designed to test this (and our account does not require it), semantic priming was evident even at the trial-by-trial level, such that seeing a category word (whether fashion, transportation, or moral) on one trial boosted recognition of subsequent category words more than it boosted recognition of subsequent noncategory words (providing such a boost of 9% for fashion, 6% for transportation, and 5% for morality [which are the means that **Gantman & Van Bavel** requested, and which straightforwardly support our account]).

R3.6.2. Really, it doesn't. Whereas some of the empirical case studies we have explored turn on subtle details that may be open to interpretation, the "moral pop-out" case study has always seemed to us to be clear, unsubtle, and unusually decisive (and we have been pleased to see that others concur; e.g., Jussim et al. 2016). **Gantman & Van Bavel** disagreed, with three primary counterarguments. However, their responses respectively (1) mischaracterize our challenge, (2) cannot possibly account for our results, and (3) bet on possibilities that are already known to be empirically false. We briefly elaborate on each of these challenges:

First, **Gantman & Van Bavel** write, "F&S recently claimed that semantic memory must be solely responsible for the moral pop-out effect because the moral words were more related to each other than the control words were." We made no such claim, and we don't even think this claim makes sense: The relatedness confound alone doesn't mean that it "*must* be *solely* responsible" (emphasis added); it merely means that semantic priming *could* be responsible, such that Gantman and Van Bavel's (2014) original conclusions wouldn't follow. Nevertheless, we actively tested this alternative empirically: When we ran the relevant experiments, semantic relatedness *in fact* produced analogous pop-out effects. It was our experimental results, not the confound itself, that suggested that "moral pop-out" is really just semantic priming.

Second, they complain that subjects in our pop-out studies were not "randomly assigned" to the three experiments we ran (i.e., fashion, transportation, or morality). This was certainly true, insofar as these were three separate experiments. But surely that can't by itself be somehow disqualifying. After all, if this feature prevented our studies of morality and fashion from being interpreted as analogous, then by the same criteria, no two experiments conducted at different times or in different labs could *ever* be compared for any purpose – even just to suggest, as we do, that both experiments appear to be investigating the same thing.

More generally, the manner in which this second complaint was supposed to undermine our argument was completely unelaborated. So, let's evaluate this carefully: Just how could such "nonrandom" assignment undermine our interpretation that morality plays no role in "moral pop-out"? If we were claiming *differences* between the experiments, then nonrandom assignment could be problematic in a straightforward way: Perhaps one group of subjects was more tired or stressed out (etc.), and that factor explains the difference. But in fact we suggested that there is *no* evidence of any relevant differences among the various pop-out effects. Our explanation for this apparent equivalence is that the same underlying process (semantic priming) drives all of the effects, with no evidence that morality, per se, plays any role. Can a lack of random assignment explain this apparent equivalence differently? Such an explanation would have to assume that the "true" effect with morality in our experiments was in fact much larger than for the other categories (due to the morality-specific boost) but that this particular group of subjects (i.e., members of the Yale community tested in one month rather than another) somehow deflated this previously undiscovered "super-pop-out" down to … exactly the same magnitude (of 4%) that Gantman & Van Bavel (2014) previously reported. In other words, "random assignment" is a red herring here, and it cannot save the day for "moral pop-out."

Third, **Gantman & Van Bavel** offer a final speculation about semantic priming to salvage their account, but in fact this speculation is demonstrably false. In particular, they suggest that semantic priming cannot explain moral pop-out because moral words cannot easily prime each other:

> We suspect that moral words are not explicitly encoded in semantic memory as moral terms or as having significant overlapping content. For example, *kill* and *just* both concern morality, but one is a noun referring to a violent act and the other is an adjective referring to an abstract property. Category priming is more likely when the terms are explicitly identifiable as being in the same category or at least as having multiple overlapping semantic features (e.g., pilot, airport). (Gantman & Van Bavel, para. 8)

But this novel suggestion completely misconstrues the nature of spreading activation in memory. Semantic priming is a phenomenon of *relatedness* – not of being "explicitly identifiable as being in the same category" – and it works just fine between nouns and adjectives (though *kill*, of course, is more commonly a verb, not a noun). Our own fashion words, for example, included words from multiple parts of speech and varying levels of abstractness (e.g., *wear, trendy, pajamas*), and they had no difficulty priming each other. And the moral words included *justice, law, illegal, crime, convict, guilty, jail*, and so on – words so related as to practically constitute a train of thought. In short, **Gantman & Van Bavel**'s speculation in this domain effectively requires that *law* and *illegal* would not activate each other via associative links in semantic memory, but this seems counter to everything we know about how semantic priming works.

R3.6.3. Labels and squiggles. Applying "labels" to meaningless squiggles (i.e., thinking of ⊓⊔ and ⊔⊓ as a rotated 2 and 5) makes them easier to find in a search array (Lupyan & Spivey 2008). Is this a "conceptual effect on visual processing" (Lupyan 2012)? Or does thinking of the symbols as familiar items just make it easier to *remember* what you're looking for? Klemfuss et al. (2012) – highlighted as one of our case studies – demonstrated the latter: When the task is repeated with a copy of the target symbol on-screen (so that one needn't remember it), the "labeling" advantage disappears; moreover, such labeling fails to improve visual processing of other features of the symbols that don't rely on memory (e.g., line thickness).

Lupyan agreed that the on-screen cue eliminated the labeling advantage but also noted that it slowed performance relative to the no-cue condition. This is simply irrelevant: The cue display was more crowded initially, and it included a stronger orienting signal (a large cue vs. a small fixation cross), both of which may have affected performance. What matters is the interaction: *holding fixed* the presence of the cue, labels had no effect, contra Lupyan's account.

More generally, though, **Lupyan**'s suggestion that our memory explanation and his "retuning of visual feature detectors" explanation (Lupyan & Spivey 2008; Lupyan et al. 2010; Lupyan & Ward 2013) are "exactly the same" is oddly self-undermining. If these effects really are explained by well-known mechanisms of memory as we suggested (see also Chen & Proctor 2012), then none of these new experiments needed to be done in the first place, because semantic priming has all of the same effects and has been well characterized for nearly half a century. By contrast, we think Lupyan's exciting and provocative work raises the revolutionary possibility that meaningfulness per se reaches down into visual processing to change what we see; but if this revolution is to be achieved, mere effects of memory must be ruled out.

R4. Whac-a-Mole

We find the prospect of a genuine top-down effect of cognition on perception to be exhilarating. In laying out our checklist of pitfalls, our genuine hope is to discover a phenomenon that survives them – and indeed many commentators suggested they had found one. On the one hand, we are hesitant to merely discuss (rather than empirically investigate) these cases, for the same reason that our target article focused so exclusively on empirical case studies: We sincerely wish to avoid the specter of vague "Australian stepbrothers" (Bruner & Goodman 1947; see sect. 5.1) that merely *could* explain away these effects, without evidence that they really do. What we really need are new empirical case studies (and we have plenty more in the works; e.g., see Firestone & Scholl 2015c). On the other hand, we have strong opinions about many of the cases raised in the commentaries – and it wouldn't be sporting to ignore them. So here, we'll discuss several of the most provocative, most compelling, best-supported cases that were raised.

In general, this part of the conversation feels a bit like the children's game of "Whac-a-Mole" (see Fig. R4): Even if you manage to whack one top-down effect, another immediately pops up to replace it. Our hope is that, by highlighting the six-pitfall checklist, such mole-whacking may occur preemptively, such that "only research reports that pass (or at least explicitly address) F&S's six criteria can henceforth become part of the serious theoretical conversation" (**Cutler & Norris**). For now, we'll play Whac-a-Mole – both for general phenomena (sect. R4.1) and specific studies (sect. R4.2).

R4.1. General phenomena

Over and above particular studies that some commentators believed escape our pitfalls, many commentaries focused on general psychological phenomena that may or may not be top-down effects of cognition on perception.

R4.1.1. Inattentional blindness and emotion-induced blindness. **Most** suggested that failures to see what is right in front of us when our attention is otherwise occupied (by a distracting task or an emotional image) are examples of cognition penetrating perception (Most et al. 2001; 2005b; Most & Wang 2011). We think these phenomena are fascinating – so much so that we wish we studied them ourselves. (Well, one of us [CF] wishes that; the other [BJS] *does* work on this topic [e.g., Ward & Scholl 2015] and thinks the second author of Most et al. 2005b made a valuable contribution.) At any rate, both of us think that "inattentional blindness" is aptly named: It is clearly a phenomenon of selective *attention*, occurring when attention is otherwise occupied. As such, it is exactly the sort of input-level effect that does not violate the encapsulation of seeing from thinking, per section R3.5.

R4.1.2. Hallucinations and delusions. In looking for top-down effects of cognition on perception, **Howe & Carter** suggested that "hallucinations are one example that clearly meets this challenge." However, **Ross, McKay, Coltheart, & Langdon** (**Ross et al.**) disagreed, arguing that two-factor theories of such abnormal psychological states "are not committed to perception being cognitively penetrable," and, indeed, "are predicated on a conceptual distinction (and empirical dissociation) between perception and cognition." We agree with Ross et al. It is important in evaluating a candidate top-down effect on perception to consider exactly what the "top" is supposed to be. If anything, hallucinations show many of the hallmarks of *in*flexible processing: After all, many patients who experience hallucinations find them to be intrusive in their daily lives and unresponsive to the patient's wishes that they would disappear.

O'Callaghan et al. suggested that hallucinations must be examples of cognitive penetrability because they incorporate autobiographical information, including visions of "familiar people or animals" such as a "deceased spouse during a period of bereavement." But this analysis conflates higher-level expectations with lower-level priming and long-term sensitivity changes; it is no coincidence, after all, that O'Callaghan et al. used "familiar" items as examples. Again, it is equally important to consider the content that hallucinations do *not* incorporate, and the states they are *not* sensitive to – including the very higher-level wishes and desires that would make these genuine top-down effects of cognition on perception.

R4.1.3. Motor expertise. **Cañal-Bruland et al.** observed that, for an unskilled baseball player facing a pitch, "the information you attune to for guiding your batting action would be crucially different from the information the expert attunes to and uses," but then they assert without argument that "this is *the* perfect example for no change in visual input but a dramatic change in visual perception" (emphasis theirs). (See also **Witt et al.**'s colorful quote by Pedro Martinez.) Why? Why is this perception at all, rather than a change in action or attentional focus? This is exactly what remains to be shown. Our core aim is to probe the distinctions between processes such as

Figure R4 (Firestone & Scholl). Some excited people playing Whac-a-mole.

perception, memory, attention, action, and so on; these are the distinctions Cañal-Bruland et al. simply ignore.

R4.1.4. Mental imagery. **Howe & Carter** suggested that, in mental imagery, "perception is obviously affected by top-down cognition" (see also **de Haas et al.** and **Esenkaya & Proulx**). We don't find this so obvious. First, Howe & Carter assert that "people actually see the mental images as opposed to merely being aware of them," without acknowledging that this is one of the single most controversial claims in the last half-century of cognitive science (for a review in this journal, see Pylyshyn 2002). But second, so what? Even if mental imagery and visual perception share some characteristics, they differ in other ways, including vividness, speed, effortfulness, and so on, and these differences allow us to distinguish visual imagery from visual perception. As **Block** argues about imagery:

> If this is cognitive penetration, why should we care about cognitive penetration? Mental imagery can be accommodated to a joint between cognition and perception by excluding these quasi-perceptual states, or alternatively, given that imagery is so slow, by excluding slow and effortful quasi-perceptual states. (Block, para. 5)

We agree.

R4.1.5. Sinewave speech. Seemingly random electronic-sounding squeaks can be suddenly and strikingly segmented into comprehensible speech when the listener first hears the unambiguous speech from which the squeaks were derived (Remez et al. 1981). **Vinson et al.** ask: "Is this not top-down?"

Maybe not. The role of auditory attention in such "sinewave speech" is still relatively unknown: Even **Vinson et al.**'s citation for this phenomenon (Darwin 1997) rejects the more robustly "top-down" interpretation of sinewave speech and instead incorporates it into a framework of "auditory grouping" – an analogy with visual grouping, which is a core phenomenon of perception and not an example of cognitive penetrability.

But maybe so. In any case, this is not our problem: As many commentaries noted, our thesis is about *visual* perception, "the most important and interesting of the human modalities" (**Keller**), and it would take a whole other manifesto to address the also-important-and-interesting case of audition. Luckily, **Cutler & Norris** have authored such a manifesto, in this very journal (Norris et al. 2000) – and their conclusion is that in speech perception, "feedback is never necessary." Bottoms up to that!

R4.1.6. Multisensory phenomena. Though the most prominent crossmodal effects are *from* vision *to* other senses (e.g., from vision to audition; McGurk & MacDonald 1976), **de Haas et al.** and **Esenkaya & Proulx** pointed to examples of other sense modalities affecting vision as evidence for cognitive penetrability. For example, a single flash of light can appear to flicker when accompanied by multiple auditory beeps (Shams et al. 2000), and waving one's hand in front of one's face while blindfolded can produce illusory motion (Dieter et al. 2014). Are these top-down effects of cognition on perception?

We find it telling that none of these empirical reports themselves connect the findings up with issues of cognitive penetrability. Indeed, these effects show the very same *inflexibility* that visual perception itself shows, and in fact they *don't* work with mere higher-level knowledge; for example, merely *knowing* that someone else is waving his or her hand in front of your face does not produce illusory motion (Dieter et al. 2014). Instead, these are straightforwardly effects of *perception* on *perception*.

R4.1.7. Drugs and "Neurosurgery." Some commentaries pointed to influences on perception from more extreme sources, including powerful hallucinogenic drugs (**Howe & Carter**) and even radical "neurosurgery" (**Goldstone et al.**). Whether raised sincerely or in jest, these cases may be exceptions that prove the rule: If the only way to get such spectacular effects on perception is to *directly alter the chemical and anatomical makeup of the brain*, then this only further testifies to the power of encapsulation and how difficult it is to observe such effects in healthy, lucid, un-operated-on observers.

R4.2. Particular studies

Beyond general phenomena that may bear on the relationship between seeing and thinking, some commentaries emphasized particular studies that they felt escaped our six pitfalls.

R4.2.1. Action-specific perception in *Pong*. Subjects judge a ball to be moving faster when playing the game *Pong* with a smaller (and thus less effective) paddle (e.g., Witt & Sugovic 2010). **Witt et al.** advertised this effect as "An action-specific effect on perception that avoids all pitfalls." We admire many of the measures this work has taken to address alternative explanations, but it remains striking that the work has still failed to apply the lessons from research on task demands. To our knowledge, subjects in these studies have never even been asked about their hypotheses (let alone told a cover story), nor have they been asked how they make their judgments. This could really matter: For example, other work on action-specific perception in similarly competitive tasks has shown that subjects blame the equipment to justify poor performance (Wesp et al. 2004; Wesp & Gasper 2012); could something similar be occurring in the *Pong* paradigm, such that subjects say the ball is moving faster to excuse their inability to catch it?

R4.2.2. Perceptual learning and object perception. **Emberson** explored a study of perceptual learning and object segmentation showing that subjects who see a target object in different orientations within a scene during a training session are subsequently more likely to see an ambiguous instance of the target object as completed behind an occluder rather than as two disconnected objects (Emberson & Amso 2012).

We find this result fascinating, but we fail to see its connection to cognitive (im)penetrability. (And we also note in passing that almost nobody in the rich field of perceptual learning has discussed their results in terms of cognitive penetrability.) **Emberson** quotes our statement that in perceptual learning, "the would-be penetrator is just the low-level input itself," but then seems to interpret this statement as referring to "simple repeated exposure." But, as we wrote right after this quoted sentence, "the thesis of cognitive impenetrability constrains the information modules can access, but it does not constrain what modules can do with the input they *do* receive." Indeed, we suspect that we have the same rich view of perceptual learning as Emberson does, such that perceptual learning may incorporate all sorts of sophisticated processing in extracting the statistics of the environment. Nevertheless, the "top" in this putative top-down effect is simply the statistical regularities of the environment.

R4.2.3. Energy and slant perception. **Clore & Proffitt** suggested that recent studies of energy and slant perception overcome demand characteristics in past research, pointing to studies of sugary beverages and estimated hill slant (Schnall et al. 2010), and quasi-experimental designs linking body weight with slant estimates of a staircase (Taylor-Covill & Eves 2016). Our target article already discussed studies of sugar and slant, in which subjects who drank a sugary beverage judged a hill to be less steep (Schnall et al. 2010). For reasons that remain unclear, all subjects in those studies also wore heavy backpacks, regardless of whether they drank a sugary beverage, and we suggested that the sugar manipulation may have interacted with the demand from the backpack. Clore & Proffitt wrote, "we are not aware of any data supporting glucose effects on susceptibility to demand." But their own commentary cited those very data: Durgin et al. (2012) empirically *demonstrated* this by showing that instructing subjects to ignore the backpack not only eliminates the backpack's effect on slant estimates (which is not so surprising), but also eliminates the effects of the *sugar* manipulation – which is quite surprising indeed, if one thinks that sugar affects perceived slant all on its own.

In another study **Clore & Proffitt** discussed, subjects were recruited at a train station and were visually classified as overweight (i.e., having a high body-mass index [BMI]) or healthy (i.e., having a normal BMI). Overweight subjects estimated a staircase as steeper (Taylor-Covill & Eves 2016), even though there was no overt manipulation to create experimental demand. However, this quasi-experimental design ensured nonrandom assignment of subjects (in a way that actually matters, due to a claimed *difference*; cf. **Gantman & Van Bavel**), and data about subjects' height and posture (etc.) were not reported, even though such variables correlate with BMI (Garn et al. 1986) and may alter subjects' staircase-viewing perspective. But more broadly, it's not clear which direction of this effect supports the view that effort affects perception. The purpose of stairs, after all, is to *decouple* steepness from effort, and in fact steeper staircases are not always harder to climb than shallower staircases, holding fixed the staircase's height of ascent. Indeed, the 23.4° staircase used in Taylor-Covill and Eves' (2016) study is actually *less steep* than the energetically optimal staircase steepness of 30° (Warren 1984), meaning that, if anything, perceiving the staircase as steeper (as the high-BMI subjects did) is actually perceiving it as *easier* to climb, not harder to climb. In other words, this effect is in the wrong direction for Clore & Proffitt's account!

R4.2.4. Categorization and inattentional blindness. **Most** reviewed evidence that categorization of a stimulus (e.g., as a number or a letter) can change the likelihood that we will see it in the first place (Most 2013). But this study manipulated categorization by changing the way the stimulus itself looked (in particular, its orientation) – the kind of low-level difference (Pitfall 4) that can really matter. Better to use a truly ambiguous stimulus (such as the B/13 stimulus employed in other top-down effects; e.g., Balcetis & Dunning 2006).

R4.2.5. Memory color. We are very impressed by **Witzel, Olkkonen, & Gegenfurtner**'s (**Witzel et al.**'s) reports that the remembered color of an object alters its perceived color – such that, for example, subjects who must set a banana's color to be gray in fact make it a bit blue (Hansen et al. 2006). This work (unlike most recently alleged top-down effects) comes from our own field and applies the rigor of vision science in careful ways that are sensitive to many of our concerns. So, what explains it?

Even if gray bananas truly look yellow, that needn't imply cognitive impenetrability; it could instead simply be a form of perceptual learning, as we explored earlier (see also Deroy 2013). Still, deep puzzles remain about the nature of this effect. For example, Hansen et al. (2006) themselves note that these memory-color effects are many times larger than established discrimination thresholds, and yet the effect fails to work as a subjectively appreciable demo: A gray banana all on its own just doesn't look yellow, and it certainly does not look as yellow as the results imply. This suggests that some kind of response bias could be involved. Because the subjects' task in these experiments is to adjust the banana's color to look gray, one possibility is that subjects are overcorrecting: They see a banana, they know that bananas are yellow, and so they try to make sure that *all of the yellow is gone*, which ends up producing a slightly blue banana. Another, less skeptical, possibility is that the effect of memory color is also an effect *on* memory – of gray. In other words, the gray *standard* that subjects have in mind as their adjustment goal may change depending on the object's identity, rather than the perceived color of the object itself (see Zeimbekis 2013; though see also Macpherson 2012).

Either way, we wonder whether this effect is truly perceptual, and we are willing to "pre-register" an experiment in this response. We suspect that, after adjusting a banana to look gray (but in fact *be* blue), subjects who see this bluish banana next to (a) an objectively gray patch and (b) a patch that is objectively as blue as the bluish (but supposedly gray-looking) banana will be able to tell that the banana is the same color as the blue patch, not the gray patch. Conversely, we suspect that subjects who see an objectively gray banana next to (a) an objectively gray patch and (b) a patch that is as yellow as the magnitude of the memory color effect will be able to tell that the banana is the same color as the gray patch, not the yellow patch (as we can in **Witzel et al.**'s figure).[5] At any rate, Witzel et al.'s account makes strong predictions to the contrary in both cases.

R5. Concluding remarks

We have a strong view about the relationship between seeing and thinking. However, the purpose of this work is not to *defend* that view. Instead, it is to lay the empirical groundwork for discovering genuinely revolutionary top-down effects of cognition on perception, which we find to be a truly exhilarating possibility. We hope we have highlighted the key empirical tools that will matter most for discovering and evaluating such effects, and that we have further shown how these tools can be employed in concrete experiments. We contend that no study has yet escaped our pitfalls; but if this framework helps reveal a genuine top-down effect of cognition on perception, we will be thrilled to have played a part.

NOTES

1. Witt et al. open their commentary with an excellent and apt sports anecdote. So here is a sports-related analogy in return: If you are concerned that a tennis line judge is biasing his or her verbal reports of whether a ball was in (perhaps because you think he or she wagered on the match), simply asking the line judge to walk over and *point* to where the ball landed will not magically remove any bias.

2. These issues have bedeviled such research for many decades, and Orne (1962) is incredibly illuminating in this respect. We thank Johan Wagemans for reminding us of its relevance and utility in this context.

3. Balcetis & Cole suggest that Durgin et al.'s (2011a) results may not be informative here because they failed to replicate the original Balcetis and Dunning (2010) beanbag result. But, in that case, the beanbag result is undermined either way: either it is explained by a strategic difference, or there is no result to explain after all (because it is not reliable). Both outcomes, of course, entail that this is not a top-down effect on perception.

4. $p < 0.05$. (For the technical details of this analysis, see Gigerenzer 2004.)

5. Note that **Witzel et al.**'s figure does not reproduce the conditions of their experiment. If the banana on the left looks yellower than the banana on the right, that's because it *is* yellower (in that it contains less blue). We find that if one "zooms in" on each banana on its own, the gray one looks gray and the blue one looks blue.

References

[The letters "a" and "r" before author's initials stand for target article and response references, respectively]

Abrams, J., Barbot, A. & Carrasco, M. (2010) Voluntary attention increases perceived spatial frequency. *Attention, Perception and Psychophysics* 72(6):1510–21. doi:10.3758/APP.72.6.1510. [BdH]

Abrams, R. A. & Weidler, B. J. (2015) How far away is that? It depends on you: Perception accounts for the abilities of others. *Journal of Experimental Psychology: Human Perception and Performance* 41:904–908. [aCF]

Adams, R. B., Gordon, H. L., Baird, A. A., Ambady, N. & Kleck, R. E. (2003) Effects of gaze on amygdala sensitivity to anger and fear faces. *Science* 300 (5625):1536. [CO]

Adams, R. B., Hess, U. & Kleck, R. E. (2015) The intersection of gender-related facial appearance and facial displays of emotion. *Emotion Review* 7(1):5–13. [CO]

Adams, R. B. & Kleck, R. E. (2005) Effects of direct and averted gaze on the perception of facially communicated emotion. *Emotion* 5(1):3. [CO]

Adelson, E. H. (2000) Lightness perception and lightness illusions. In: *The new cognitive neurosciences, second edition*, ed. M. Gazzaniga, pp. 339–51. MIT Press. [aCF]

Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C. & de Lange, F. P. (2013) Shared representations for working memory and mental imagery in early visual cortex. *Current Biology* 23:1427–31. [PDLH]

Alink, A., Schwiedrzik, C. M., Kohler, A., Singer, W. & Muckli, L. (2010) Stimulus predictability reduces responses in primary visual cortex. *The Journal of Neuroscience* 30(8):2960–66. [LMH]

Alter, A. L. & Balcetis, E. (2011) Fondness makes the distance grow shorter: Desired locations seem closer because they seem more vivid. *Journal of Experimental Social Psychology* 47:16–21. [aCF]

Amaral, D. G., Behniea, H. & Kelly, J. L. (2003) Topographic organization of projections from the amygdala to the visual cortex in the macaque monkey. *Neuroscience* 118(4):1099–120. [DWV]

Aminoff, E., Gronau, N. & Bar, M. (2007) The parahippocampal cortex mediates spatial and nonspatial associations. *Cerebral Cortex* 17(7):1493–503. [CO]

Anderson, A. K. & Phelps, E. A. (2001) Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature* 411:305–309. [APG]

Anderson, M. L. (2014) *After phrenology: Neural reuse and the interactive brain.* MIT Press. [LMH]

Angelucci, A., Levitt, J. B., Walton, E. J., Hupe, J.-M., Bullier, J. & Lund, J. S. (2002) Circuits for local and global signal integration in primary visual cortex. *The Journal of Neuroscience* 22(19):8633–46. [CO]

Anllo-Vento, L., Luck, S. J. & Hillyard, S. A. (1998) Spatio-temporal dynamics of attention to color: Evidence from human electrophysiology. *Human Brain Mapping* 6:216–38. [AR]

Anstis, S. (2002) Was El Greco astigmatic? *Leonardo* 35:208. [aCF]

Anton-Erxleben, K., Abrams, J. & Carrasco, M. (2011) Equality judgments cannot distinguish between attention effects on appearance and criterion: A reply to Schneider (2011) *Journal of Vision* 11(13):8. doi:10.1167/11.13.8. [BdH]

Anton-Erxleben, K. & Carrasco, M. (2013) Attentional enhancement of spatial resolution: Linking behavioural and neurophysiological evidence. *Nature Reviews Neuroscience* 14(3):188–200. doi:10.1038/nrn3443. [BdH, DWV]

Anton-Erxleben, K., Henrich, C. & Treue, S. (2007) Attention changes perceived size of moving visual patterns. *Journal of Vision* 7(5):1–9. [aCF]

Bach-y-Rita, P. & Kercel, S. W. (2003) Sensory substitution and the human–machine interface. *Trends in Cognitive Sciences* 7(12):541–46. doi:S1364661303002900 [pii]. [TE]

Baker, L. J. & Levin, D. T. (2016) The face-race lightness illusion is not driven by low-level stimulus properties: An empirical reply to Firestone & Scholl (2014). *Psychonomic Bulletin and Review*. doi:10.3758/s13423-016-1048-z [DTL]

Balcetis, E. (2016) Approach and avoidance as organizing structures for motivated distance perception. *Emotion Review* 8:115–28. [aCF]

Balcetis, E. & Dunning, D. (2006) See what you want to see: Motivational influences on visual perception. *Journal of Personality and Social Psychology* 91:612–25. [arCF]

Balcetis, E. & Dunning, D. (2010) Wishful seeing: More desired objects are seen as closer. *Psychological Science* 21:147–52. [arCF, EB]

Balcetis, E., Dunning, D. & Granot, Y. (2012) Subjective value determines initial dominance in binocular rivalry. *Journal of Experimental Social Psychology* 48:122–29. [EB, aCF]

Baldauf, D. & Desimone, R. (2014) Neural mechanisms of object-based attention. *Science* 344(6182):424–27. Availabile at: http://doi.org/10.1126/science.1247003. [DMB]

Banerjee, P., Chatterjee, P. & Sinha, J. (2012) Is it light or dark? Recalling moral behavior changes perception of brightness. *Psychological Science* 23:407–409. [PT, aCF]

Bannert, M. M. & Bartels, A. (2013) Decoding the yellow of a gray banana. *Current Biology* 23(22):2268–72. [CW, aCF]

Bar, M. (2004) Visual objects in context. *Nature Reviews Neuroscience* 5(8):617–29. [CO]

Bar, M. (2007) The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences* 11(7):280–89. [LMH]

Bar, M. & Aminoff, E. (2003) Cortical analysis of visual context. *Neuron* 38(2):347–58. [CO]

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., Hämäläinen, M., Marinkovic, K., Schacter, D. & Rosen, B. (2006) Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences USA* 103(2):449–54. [CO]

Bar, M., Tootell, R. B., Schacter, D. L., Greve, D. N., Fischl, B., Mendola, J. D., Rosen, B. R. & Dale, A. M. (2001) Cortical mechanisms specific to explicit visual object recognition. *Neuron* 29(2):529–35. [CO]

Barbas, H. (2015) General cortical and special prefrontal connections: Principles from structure to function. *Annu Rev Neurosci* 38:269–89. [LMH]

Bargh, J. A. & Chartrand, T. L. (2000) The mind in the middle: A practical guide to priming and automaticity research. In: *Handbook of research methods in social and personality psychology, second edition*, ed. H. T. Reis & C. M. Judd, pp. 253–85. Cambridge University Press. [GLC]

Barnes, J. & David, A. S. (2001) Visual hallucinations in Parkinson's disease: A review and phenomenological survey. *Journal of Neurology, Neurosurgery, and Psychiatry* 70(6):727–33. doi:10.1136/jnnp.70.6.727. [CO]

Barr, M. (2009) The proactive brain: Memory for predictions. *Philosophical Transactions of the Royal Society London B Biological Sciences* 364:1235–43. [AR]

Barrett, H. & Kurzban, R. (2006) Modularity in cognition. *Psychological Review* 113:628–47. [RO]

Barrett, L. F. (2009) The future of psychology: Connecting mind to brain. *Perspectives on Psychological Science* 4(4):326–39. [LMH]

Barrett, L. F. & Satpute, A. B. (2013) Large-scale brain networks in affective and social neuroscience: Towards an integrative functional architecture of the brain. *Current Opinion in Neurobiology* 23(3):361–72. [LMH]

Barrett, L. F. & Simmons, W. K. (2015) Interoceptive predictions in the brain. *Nature Reviews Neuroscience* 16:419–29. [LMH]

Barrett, L. F., Tugade, M. M. & Engle, R. W. (2004) Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin* 130:553–73. PMCID: PMC1351135. [LMH]

Barsalou, L. W. (2010) Grounded cognition: Past, present, and future. *Topics in Cognitive Science* 2:716–24. [WG]

Baruch, O. & Yeshurun, Y. (2014) Attentional attraction of receptive fields can explain spatial and temporal effects of attention. *Visual Cognition* 22(5):704–36. doi:10.1080/13506285.2014.911235. [BdH]

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P. & Friston, K. J. (2012) Canonical microcircuits for predictive coding. *Neuron* 76:695–711. [ACl]

Beale, J. M. & Keil, F. C. (1995) Categorical effects in the perception of faces. *Cognition* 57:217–39. [WG]

Behrens, T. E. J., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. A. (2007) Learning the value of information in an uncertain world. *Nature Neuroscience* 10:1214–21. [LMH]

Berger, T. D., Martelli, M. & Pelli, D. G. (2003) Flicker flutter: Is an illusory event as good as the real thing? *Journal of Vision* 3(6):406–12. doi:10:1167/3.6.1. [BdH]

Berthenhal, B. I., Campos, J. J. & Haith, M. M. (1980) Development of visual organization: The perception of subjective contours. *Child Development* 51:1072–80. [GSH]

Bhalla, M. & Proffitt, D. R. (1999) Visual-motor recalibration in geographical slant perception. *Journal of Experimental Psychology: Human Perception and Performance* 25:1076–96. [arCF, FHD]

Bieszczad, K. M. & Weinberger, N. M. (2010) Representational gain in cortical area underlies increase of memory strength. *Proceedings of the National Academy of Sciences USA* 107:3793–98. [VM]

Blaesi, S. & Bridgeman, B. (2015) Perceived difficulty of a motor task affects memory but not action. *Attention, Perception, and Psychophysics* 77(3):972–77. [PT]

Blake, R. & Sekuler, R. (2005) *Perception, fifth edition*. McGraw-Hill. [aCF]

Blasdel, G. G. & Salama, G. (1986) Voltage-sensitive dyes reveal a modular organization in monkey striate cortex. *Nature* 321:579–85. [PDLH]

Boothe, R. G., Dobson, V. & Teller, D. Y. (1985) Postnatal development of vision in human and nonhuman primates. *Annual Review of Neuroscience* 8:495–545. [GSH]

Boutonnet, B. & Lupyan, G. (2015) Words jump-start vision: A label advantage in object recognition. *Journal of Neuroscience* 32(25):9329–35. doi:10.1523/jneurosci.5111-14.2015. [MR, GL]

Bouvier, S. & Treisman, A. (2010) Visual feature binding requires reentry. *Psychological Science* 21:200–204. [aCF]

Briggs, F. & Usrey, W. M. (2005) Temporal properties of feedforward and feedback pathways between the thalamus and visual cortex in the ferret. *Thalamus and Related Systems* 3:133–39. [VM]

Brighetti, G., Bonifacci, P., Borlimi, R. & Ottaviani, C. (2007) "Far from the heart far from the eye": Evidence from the Capgras delusion. *Cognitive Neuropsychiatry* 12(3):189–97. [RMR]

Brockmole, J. R., Wang, R. F. & Irwin, D. E. (2002) Temporal integration between visual images and visual percepts. *Journal of Experimental Psychology: Human Perception and Performance* 28(2):315–34. [NB]

Bruce, V., Langton, S. & Hill, H. (1999) Complexities of face perception and categorisation. *Behavioral and Brain Sciences* 22(3):369–70. [DWV]

Bruner, J. S. (1957) On perceptual readiness. *Psychological Review* 64:123–52. [aCF]

Bruner, J. S. & Goodman, C. C. (1947) Value and need as organizing factors in perception. *Journal of Abnormal and Social Psychology* 42:33–44. [arCF]

Bruner, J. S., Postman, L. & Rodrigues, J. (1951) Expectation and the perception of color. *The American Journal of Psychology* 64(2):216–27. [CW, aCF]

Bruns, P., Maiworm, M. & Röder, B. (2014) Reward expectation influences audiovisual spatial integration. *Attention, Perception and Psychophysics* 76(6):1815–27. doi:10.3758/s13414-014-0699-y. [BdH]

Bubl, E., Kern, E., Ebert, D., Bach, M. & Tebartz van Elst, L. (2010) Seeing gray when feeling blue? Depression can be measured in the eye of the diseased. *Biological Psychiatry* 68:205–208. [aCF]

Buffalo, E. A., Bertini, G., Ungerleider, L. G. & Desimone, R. (2005) Impaired filtering of distracter stimuli by TE neurons following V4 and TEO lesions in Macaques. *Cerebral Cortex* 15(2):141–51. Available at: http://doi.org/10.1093/cercor/bhh117. [DMB]

Bullier, J. (1999) Visual perception is too fast to be impenetrable to cognition. *Behavioral and Brain Sciences* 22(3):370. [DWV]

Bullier, J. (2001) Integrated model of visual processing. *Brain Research Reviews* 36 (2):96–107. [CO]

Bulthoff, I., Bulthoff, H. & Sinha, P. (1998) Top-down influences on stereoscopic depth-perception. *Nature Neuroscience* 1(3):254–57. [GL]

Burns, B. & Shepp, B. E. (1988) Dimensional interactions and the structure of psychological space: The representation of hue, saturation, and brightness. *Perception and Psychophysics* 43:494–507. [RG]

Cañal-Bruland, R., Pijpers, J. R. R. & Oudejans, R. R. D. (2010) The influence of anxiety on action-specific perception. *Anxiety, Stress and Coping* 23:353–61. [aCF]

Cañal-Bruland, R. & van der Kamp, J. (2015) Embodied perception: A proposal to reconcile affordance and spatial perception. i-Perception 6:63–66. [RC-B]

Cañal-Bruland, R., Zhu, F. F., van der Kamp, J. & Masters, R. S. W. (2011) Target-directed visual attention is a prerequisite for action-specific perception. *Acta Psychologica* 136:285–89. [aCF]

Carandini, M. (2012) From circuits to behavior: A bridge too far? *Nature Neuroscience* 15:507–509. [rCF]

Carandini, M., Demb, J. B., Mante, V., Tolhurst, D. J., Dan, Y., Olshausen, B. A., Gallant, J.L. & Rust, N.C. (2005) Do we understand what the early visual system does? *The Journal of Neuroscience* 25:10577–97. [VM]

Carmel, D., Thorne, J. D., Rees, G. & Lavie, N. (2011) Perceptual load alters visual excitability. *Journal of Experimental Psychology: Human Perception and Performance* 37(5):1350–60. doi:10.1037/a0024320. [BdH]

Carrasco, M. (2011) Visual attention: The past 25 years. *Vision Research* 51:1484–525. [rCF, NB, DWV]

Carrasco, M., Fuller, S. & Ling, S. (2008) Transient attention does increase perceived contrast of suprathreshold stimuli: A reply to Prinzmetal, Long, and Leonhardt (2008) *Perception and Psychophysics* 70(7):1151–64. doi:10.3758/PP.70.7.1151. [BdH]

Carrasco, M., Ling, S. & Read, S. (2004) Attention alters appearance. *Nature Neuroscience* 7:308–13. [ACl, aCF]

Carruthers, P. (2006) *The architecture of the mind: Massive modularity and the flexibility of thought*. tOxford University Press. [RO]

Carter, L. F. & Schooler, K. (1949) Value, need, and other factors in perception. *Psychological Review* 56:200–207. [aCF]

Caruso, E. M., Mead, N. L. & Balcetis, E. (2009) Political partisanship influences perception of biracial candidates' skin tone. *Proceedings of the National Academy of Sciences USA* 106:20168–73. [aCF]

Cavanagh, P. (1999) The cognitive penetrability of cognition. *Behavioral and Brain Sciences* 22(3):370–71. [DWV]

Cave, K. R. & Bichot, N. P. (1999) Visuospatial attention: Beyond a spotlight model. *Psychonomic Bulletin and Review* 6:204–23. [aCF]

Chanes, L. & Barrett, L. F. (2016) Redefining the role of limbic areas in cortical processing. *Trends in Cognitive Sciences* 20(2):96–106. [LMH]

Changizi, M. A. & Hall, W. G. (2001) Thirst modulates a perception. *Perception* 30:1489–97. [aCF]

Chaumon, M., Kveraga, K., Barrett, L. F. & Bar, M. (2013) Visual predictions in the orbitofrontal cortex rely on associative content. *Cerebral Cortex* 24(11):2899–907. doi:10.1093/cercor/bht146. [CO]

Chelazzi, L., Miller, E., Duncan, J. & Desimone, R. (1993) A neural basis for visual search in inferior temporal cortex. *Nature* 363:345–47. [AR]

Chen, H. & Scholl, B. J. (2013) Congruence with items held in visual working memory boosts invisible stimuli into awareness: Evidence from motion-induced blindness [Abstract]. *Journal of Vision* 13(9):808. [aCF]

Chen, J. & Proctor, R. W. (2012) Influence of category identity on letter matching: Conceptual penetration of visual processing or response competition? *Attention, Perception, and Psychophysics* 74:716–29. [rCF]

Chen, Y.-C. & Scholl, B. J. (2014) Seeing and liking: Biased perception of ambiguous figures consistent with the "inward bias" in aesthetic preferences. *Psychonomic Bulletin and Review* 21:1444–51. [rCF]

Chen, Y.-C. & Scholl, B. J. (2016) The perception of history: Seeing causal history in static shapes induces illusory motion perception. *Psychological Science* 27:923–30. [rCF, GL]

Chikazoe, J., Lee, D. H., Kriegeskorte, N. & Anderson, A. K. (2014) Population coding of affect across stimuli, modalities and individuals. *Nature Neuroscience* 17:1114–22. [VM]

Chua, K. W. & Gauthier, I. (2015) Learned attention in an object-based frame of reference. *Journal of Vision* 15(12):899–99. [DWV]

Chun, M. M. & Potter, M. C. (1995) A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance* 21(1):109–27. [SBM]

Chung, S. T. & Pease, P. L. (1999) Effect of yellow filters on pupil size. *Optometry and Vision Science* 76(1):59–62. [WX]

Churchland, P. M. (1988) Perceptual plasticity and theoretical neutrality: A reply to Jerry Fodor. *Philosophy of Science* 55:167–87. [aCF]

Churchland, P. S., Ramachandran, V. S. & Sejnowski, T. J. (1994) A critique of pure vision. In: *Large-scale neuronal theories of the brain*, ed. C. Koch & J. Davis, pp. 23–60. MIT Press. [aCF]

Cisek, P. & Klaska, J. F. (2010) Neural mechanisms for interacting with a world full of action choices. *Annual Review of Neuroscience* 33:269–98. [LMH]

Clark, A. (2013) Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences* 36(3):181–204. [ACl, LMH, AR, aCF]

Clark, A. (2016) *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford University Press. [ACl]

Clark-Polner, E., Wager, T. D., Satpute, A. B. & Barrett, L. F. (2016) "Brain-based fingerprints?" Variation, and the search for neural essences in the science of emotion. In: *The handbook of emotion, fourth edition*, ed. L. F. Barrett, M. Lewis & J. M. Haviland-Jones, pp. 146–65. Guilford. [LMH]

Clavagnier, S., Falchier, A. & Kennedy, H. (2004) Long-distance feedback projections to area V1: Implications for multisensory integration, spatial awareness, and visual consciousness. *Cognitive, Affective, and Behavioral Neuroscience* 4 (2):117–26. [DWV]

Clerkin, E. M., Cody, M. W., Stefanucci, J. K., Proffitt, D. R. & Teachman, B. A. (2009) Imagery and fear influence height perception. *Journal of Anxiety Disorders* 23:381–86. [aCF]

Cohen, M. A., Alvarez, G. A. & Nakayama, K. (2011) Natural-scene perception requires attention. *Psychological Science* 22:1165–72. [aCF]

Cole, S. & Balcetis, E. (2013) Sources of resources: Bioenergetic and psychoenergetic resources influence distance perception. *Social Cognition* 31:721–32. [EB, rCF]

Cole, S., Balcetis, E. & Dunning, D. (2012) Affective signals of threat increase perceived proximity. *Psychological Science* 24:34–40. [aCF]

Cole, S., Balcetis, E. & Zhang, S. (2013) Visual perception and regulatory conflict: Motivation and physiology influence distance perception. *Journal of Experimental Psychology: General* 142:18–22. [aCF]

Collins, A. M. & Loftus, E. F. (1975) A spreading-activation theory of semantic processing. *Psychological Review* 82:407–28. [aCF]

Collins, J. A. & Olson, I. R. (2014) Knowledge is power: How conceptual knowledge transforms visual cognition. *Psychonomic Bulletin and Review* 21:843–60. [PDLH, aCF]

Coltheart, M. (2007) The 33rd Sir Frederick Bartlett lecture: Cognitive neuropsychiatry and delusional belief. *The Quarterly Journal of Experimental Psychology* 60(8):1041–62. [RMR]

Coltheart, M. (2013) On the distinction between monothematic and polythematic delusions. *Mind and Language* 28(1):103–12. [RMR]

Coltheart, M., Langdon, R. & McKay, R. (2011) Delusional belief. *Annual Review of Psychology* 62(5):271–98. [RMR]

Conway, C., Jones, B., DeBruine, L., Welling, L., Smith, M. L., Perrett, D., Sharp, M. A. & Al-Dujaili, E. A. (2007) Salience of emotional displays of danger and contagion in faces is enhanced when progesterone levels are raised. *Hormones and Behavior* 51(2):202–206. [CO]

Cooper, A. D., Sterling, C. P., Bacon, M. P. & Bridgeman, B. (2012) Does action affect perception or memory? *Vision Research* 62:235–40. [PT, aCF]

Coren, S. & Enns, J. T. (1993) Size contrast as a function of conceptual similarity between test and inducers. *Perception and Psychophysics* 54:579–88. [aCF]

Corlett, P. R., Taylor, J. R., Wang, X. J., Fletcher, P. C. & Krystal, J. H. (2010) Toward a neurobiology of delusions. *Progress in Neurobiology* 92(3):345–69. [RMR]

Correll, J., Wittenbrink, B., Crawford, M. T. & Sadler, M. S. (2015) Stereotypic vision: How stereotypes disambiguate visual stimuli. *Journal of Personality and Social Psychology* 108:219–33. [aCF]

Çukur, T., Nishimoto, S., Huth, A. G. & Gallant, J. L. (2013) Attention during natural vision warps semantic representation across the human brain. *Nature Neuroscience* 16(6):763–70. Available at: http://doi.org/10.1038/nn.3381. [DMB, GL]

D'Esposito, M. & Postle, B. R. (2015) The cognitive neuroscience of working memory. *Annual Review of Psychology* 66:115–42. Available at: http://doi.org/10.1146/annurev-psych-010814-015031. [GL]

Dambacher, M., Rolfs, M., Göllner, K., Kliegl, R. & Jacobs, A. M. (2009) Event-related potentials reveal rapid verification of predicted visual input. *PLoS ONE* 4(3):e5047. doi:10.1371/journal.pone.0005047. [MR]

Darwin, C. J. (1997) Auditory grouping. *Trends in Cognitive Sciences* 1(9):327–33. [DWV, rCF]

David, S. V., Vinje, W. E. & Gallant, J. L. (2004) Natural stimulus statistics alter the receptive field structure of V1 neurons. *The Journal of Neuroscience* 24 (31):6991–7006. [DWV]

Dayan, P., Hinton, G. E., Neal, R. & Zemel, R. (1995) The Helmholtz machine. *Neural Computation* 7(5):889–904. [DWV]

de Gelder, B. & Tamietto, M. (2011) Faces, bodies, social vision as agent vision, and social consciousness. In: *The science of social vision*, ed. R. B. Adams, N. Ambady, K. Nakayama & S. Shimojo, pp. 51–74. Oxford University Press. [CO]

de Haas, B., Cecere, R., Cullen, H., Driver, J. & Romei, V. (2013a) The duration of a co-occurring sound modulates visual detection performance in humans. *PLoS ONE* 8(1):e54789. doi:10.1371/journal.pone.0054789. [BdH]

de Haas, B., Kanai, R., Jalkanen, L. & Rees, G. (2012) Grey matter volume in early human visual cortex predicts proneness to the sound-induced flash illusion. *Proceedings of the Royal Society B: Biological Sciences* 279(1749):4955–61. doi:10.1098/rspb.2012.2132. [BdH]

de Haas, B. & Rees, G. (2010) Multiple stages of cross-modal integration in visual processing: Comment on "Crossmodal influences on visual perception" by L. Shams & R. Kim. *Physics of Life Reviews* 7:287–88; discussion 29. PMID 20598657 DOI: 10.1016/j.plrev.2010.06.007. [BdH]

de Haas, B., Schwarzkopf, D. S., Anderson, E. J. & Rees, G. (2014) Perceptual load affects spatial tuning of neuronal populations in human early visual cortex. *Current Biology*: 24(2):R66–67. doi:10.1016/j.cub.2013.11.061. [BdH]

de Haas, B., Schwarzkopf, D. S., Urner, M. & Rees, G. (2013b) Auditory modulation of visual stimulus encoding in human retinotopic cortex. *NeuroImage* 70:258–67. doi:10.1016/j.neuroimage.2012.12.061. [BdH]

Delk, J. & Fillenbaum, S. (1965) Differences in perceived color as a function of characteristic color. *The American Journal of Psychology* 78:290–93. [NB]

den Daas, C., Häfner, M. & de Wit, J. (2013) Sizing opportunity: Biases in estimates of goal-relevant objects depend on goal congruence. *Social Psychological and Personality Science* 4:362–68. [aCF]

den Ouden, H. E. M., Daunizeau, J., Roiser, J., Friston, K. & Stephan, K. (2010) Striatal prediction error modulates cortical coupling. *Journal of Neuroscience* 30:3210–19. [ACl]

Den Ouden, H. E., Kok, P. & De Lange, F. P. (2012) How prediction errors shape perception, attention, and motivation. *Frontiers in Psychology* 3:548. [LMH]

Deroy, O. (2013) Object-sensitivity versus cognitive penetrability of perception. *Philosophical Studies* 162:87–107. [rCF]

Descartes, R. (1989) *On the passions of the soul*. Hackett (Original work published 1649). [LMH]

Desseilles, M., Balteau, E., Sterpenich, V., Dang-Vu, T. T., Darsaud, A., Vandewalle, G., Albouy, G., Salmon, E., Peters, F., Schmidt, C., Schabus, M., Gais, S., Degueldre, C., Phillips, C., Luxen, A., Ansseau, M., Maquet, P. & Schwartz, S. (2009) Abnormal neural filtering of irrelevant visual information in depression. *Journal of Neuroscience* 29(5):1395–403. [MD]

Desseilles, M., Schwartz, S., Dang-Vu, T. T., Sterpenich, V., Ansseau, M., Maquet, P. & Phillips, C. (2011) Depression alters "top-down" visual attention: A dynamic causal modeling comparison between depressed and healthy subjects. *NeuroImage* 54(2):1662–68. [MD]

Dicks, M., Button, C. & Davids, K. (2010) Examination of gaze behaviors under in situ and video simulation task constraints reveals differences in information pickup for perception and action. *Attention, Perception, and Psychophysics* 72:706–20. [RC-B]

Dieter, K. C., Hu, B., Knill, D. C., Blake, R. & Tadin, D. (2014) Kinesthesis can make an invisible hand visible. *Psychological Science* 25(1):66–75. doi:10.1177/0956797613497968. [BdH, rCF]

Dils, A. T. & Boroditsky, L. (2010a) Processing unrelated language can change what you see. *Psychonomic Bulletin and Review* 17:882–88. [aCF]

Dils, A. T. & Boroditsky, L. (2010b) Visual motion aftereffect from understanding motion language. *Proceedings of the National Academy of Sciences USA* 107:16396–400. [NB, aCF]

Doerrfeld, A., Sebanz, N. & Shiffrar, M. (2012) Expecting to lift a box together makes the load look lighter. *Psychological Research* 76:467–75. [aCF]

Douglas, R. J. & Martin, K. A. C. (2007) Mapping the matrix: The ways of neocortex. *Neuron* 56:226–38. [VM]

Driver, J. & Noesselt, T. (2008) Multisensory interplay reveals crossmodal influences on "sensory-specific" brain regions, neural responses, and judgments. *Neuron* 57(1):11–23. doi:10.1016/j.neuron.2007.12.013. [BdH]

Dulany, D. E. (1962) The place of hypotheses and intentions: An analysis of verbal control in verbal conditioning. *Journal of Personality* 30:102–29. [GLC]

Dumoulin, S. O. & Wandell, B. A. (2008) Population receptive field estimates in human visual cortex. *NeuroImage* 39(2):647–60. doi:10.1016/j.neuroimage.2007.09.034. [BdH]

Duncker, K. (1939) The influence of past experience upon perceptual properties. *American Journal of Psychology* 52(2):255–65. [CW]

Dunning, D. & Balcetis, E. (2013) Wishful seeing: How preferences shape visual perception. *Current Directions in Psychological Science* 22:33–37. [aCF]

Durgin, F. H. (2014) Angular scale expansion theory and the misperception of egocentric distance in locomotor space. *Psychology and Neuroscience* 7:253–60. [FHD]

Durgin, F. H., Baird, J. A., Greenburg, M., Russell, R., Shaughnessy, K. & Waymouth, S. (2009) Who is being deceived? The experimental demands of wearing a backpack. *Psychonomic Bulletin and Review* 16:964–69. [GLC, FHD, arCF]

Durgin, F. H., Hajnal, A., Li, Z., Tonge, N. & Stigliani, A. (2010) Palm boards are not action measures: An alternative to the two-systems theory of geographical slant perception. *Acta Psychologica* 134:182–97. [FHD]

Durgin, F. H., DeWald, D., Lechich, S., Li, Z. & Ontiveros, Z. (2011a) Action and motivation: Measuring perception or strategies? *Psychonomic Bulletin and Review* 18(6):1077–82. [EB, PT, arCF]

Durgin, F. H., Hajnal, A., Li, Z., Tonge, N. & Stigliani, A. (2011b) An imputed dissociation might be an artifact: Further evidence for the generalizability of the observations of Durgin et al. 2010. *Acta Psychologica* 138:281–84. [FHD]

Durgin, F. H., Klein, B., Spiegel, A., Strawser, C. J. & Williams, M. (2012) The social psychology of perception experiments: Hills, backpacks, glucose, and the problem of generalizability. *Journal of Experimental Psychology: Human Perception and Performance* 38:1582–95. [GLC, FHD, arCF]

Eacott, M. J. & Heywood, C. A. (1995) Perception and memory: Action and interaction. *Critical Reviews in Neurobiology* 9(4):311–20. [CW]

Egner, T. & Hirsch, J. (2005) Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience* 8:1784–90. [APG]

Egner, T., Monti, J. M. & Summerfield, C. (2010) Expectation and surprise determine neural population responses in the ventral visual stream. *The Journal of Neuroscience* 30(49):16601–608. [LMH]

Ellis, H. D., Young, A. W., Quayle, A. H. & de Pauw, K. W. (1997) Reduced autonomic responses to faces in Capgras delusion. *Proceedings of the Royal Society B – Biological Sciences* 264(1384):1085–92. [RMR]

Elman, J. L. & McClelland, J. L. (1988) Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language* 27:143–65. [ACu]

Emberson, L. L. & Amso, D. (2012) Learning to sample: Eye tracking and fMRI indices of changes in object perception. *Journal of Cognitive Neuroscience* 24:2030–42. [rCF, LLE]

Erdelyi, M. (1974) A new look at the new look: Perceptual defense and vigilance. *Psychological Review* 81:1–25. [aCF]

Esterman, M. & Yantis, S. (2010) Perceptual expectation evokes category-selective cortical activity. *Cerebral Cortex* 20(5):1245–53. Available at: http://doi.org/10.1093/cercor/bhp188. [DMB]

Ewbank, M. P., Fox, E. & Calder, A. J. (2010) The interaction between gaze and facial expression in the amygdala and extended amygdala is modulated by anxiety. *Frontiers in Human Neuroscience* 4(56):1–11. [CO]

Fantoni, C. & Gerbino, W. (2014) Body actions change the appearance of facial expressions. *PLoS ONE* 9:e108211. doi:10.1371/journal.pone.0108211. [WG]

Fantoni, C., Rigutti, S. & Gerbino, W. (2016) Bodily action penetrates affective perception. *PeerJ* 4:e1677. doi:10.7717/peerj.1677. [WG]

Feldman, H. & Friston, K. J. (2010) Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience* 4:215. [ACl]

Felleman, D. J. & Van Essen, D. C. (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex* 1(1):1–47. [CO]

ffytche, D. H., Howard, R. J., Brammer, M. J., David, A., Woodruff, P. & Williams, S. (1998) The anatomy of conscious vision: An fMRI study of visual hallucinations. *Nature Neuroscience* 1:738–42. [PDLH]

Finger, S. (2001) *Origins of neuroscience: A history of explorations into brain function*. Oxford University Press. [LMH]

Firestone, C. (2013a) How "paternalistic" is spatial perception? Why wearing a heavy backpack doesn't – and *couldn't* – make hills look steeper. *Perspectives on Psychological Science* 8:455–73. [aCF]

Firestone, C. (2013b) On the origin and status of the "El Greco fallacy." *Perception* 42:672–74. [aCF]

Firestone, C. & Scholl, B. J. (2014a) "Please tap the shape, anywhere you like": Shape skeletons in human vision revealed by an exceedingly simple measure. *Psychological Science* 25(2):377–86. Available at: http://doi.org/10.1177/0956797613507584. [GL, rCF]

Firestone, C. & Scholl, B. J. (2014b) "Top-down" effects where none should be found: The El Greco fallacy in perception research. *Psychological Science* 25:38–46. [arCF, WX]

Firestone, C. & Scholl, B. J. (2015a) Can you *experience* top-down effects on perception? The case of race categories and perceived lightness. *Psychonomic Bulletin and Review* 22:694–700. [arCF, WX, DTL]

Firestone, C. & Scholl, B. J. (2015b) Enhanced visual awareness for morality and pajamas? Perception vs. memory in 'top-down' effects. *Cognition* 136:409–16. [APG, arCF]

Firestone, C. & Scholl, B. J. (2015c). When do ratings implicate perception versus judgment?: The "overgeneralization test" for top-down effects. *Visual Cognition* 23(9–10):1217–26. [rCF]

Firestone, C. & Scholl, B. J. (2016) "Moral perception" reflects neither morality nor perception. *Trends in Cognitive Sciences* 20:75–76. [rCF]

Fletcher, P. C. & Frith, C. D. (2009) Perceiving is believing: A Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews Neuroscience* 10(1):48–58. [RMR]

Fodor, J. A. (1974) Special sciences (or: The disunity of science as a working hypothesis). *Synthese* 28:97–115. [rCF]

Fodor, J. A. (1983) *Modularity of mind: An essay on faculty psychology*. MIT Press. [RO, NB, aCF, BdH, LMH, DWV]

Fodor, J. A. (1984) Observation reconsidered. *Philosophy of Science* 51:23–43. [aCF]

Fodor, J. A. (1988) A reply to Churchland's "Perceptual plasticity and theoretical neutrality." *Philosophy of Science* 55:188–98. [aCF]

Fodor, J. A. & Pylyshyn, Z. W. (1981) How direct is visual perception? Some reflections on Gibson's "ecological approach." *Cognition* 9(2):139–96. [PT]

Fodor, J. A. & Pylyshyn, Z. W. (1988) Connectionism and cognitive architecture: A critical analysis. *Cognition* 28:3–71. [GSH]

Fontanini, A. & Katz, D. B. (2008) Behavioral states, network states, and sensory response variability. *Journal of Neurophysiology* 100:1160–68. [VM]

Ford, J. M., Palzes, V. A., Roach, B. J., Potkin, S. G., van Erp, T. G., Turner, J. A., Mueller, B. A., Calhoun, V. D., Voyvodic, J., Belger, A., Bustillo, J., Vaidya, J. G., Preda, A., McEwen, S. C. (2015) Visual hallucinations are associated with hyperconnectivity between the amygdala and visual cortex in people with a diagnosis of schizophrenia. *Schizophrenia Bulletin* 41:223–32. [PDLH]

Fox, E., Mathews, A., Calder, A. J. & Yiend, J. (2007) Anxiety and sensitivity to gaze direction in emotionally expressive faces. *Emotion* 7(3):478. [CO]

Fox, R. A. (1984) Effect of lexical status on phonetic categorization. *Journal of Experimental Psychology: Human Perception and Performance* 10:526–40. [ACu]

Franconeri, S. L., Alvarez, G. A. & Cavanagh, P. (2013) Flexible cognitive resources: Competitive content maps for attention and memory. *Trends in Cognitive Sciences* 17:134–41. [VM]

Frank, M. J., Rudy, J. W. & O'Reilly, R. C. (2003) Transitivity, flexibility, conjunctive representations, and the hippocampus. II. A computational analysis. *Hippocampus* 13(3):341–54. doi:10.1002/hipo.10084. [LLE]

Freeman, J. B., Rule, N. O., Adams, R. B. & Ambady, N. (2010) The neural basis of categorical face perception: Graded representations of face gender in fusiform and orbitofrontal cortices. *Cerebral Cortex* 20(6):1314–22. [CO]

Friston K (2010) The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience* 11:127–38. [LMH, DWV]

Friston, K. J., Ashburner, J., Kiebel, S. J., Nichols, T. E. & Penny, W. D., eds. (2007) *Statistical parametric mapping: The analysis of functional brain images*. Academic Press. [MD]

Friston, K. J. & Buchel, C. (2000) Attentional modulation of effective connectivity from V2 to V5/MT in humans. *Proceedings of the National Academy of Sciences USA* 97(13):7591–96. [MD]

Friston, K. J., Harrison, L. & Penny, W. (2003) Dynamic causal modelling. *NeuroImage* 19(4):1273–302. [MD]

Frith, C. D. & Friston, K. J. (2013) False perceptions and false beliefs: Understanding schizophrenia. *Neurosciences and the Human Person: New Perspectives on Human Activities* 121:1–15. [RMR]

Gall, F. J. (1835) *On the functions of the brain and of each of its parts: With observations on the possibility of determining the instincts, propensities, and talents, or the moral and intellectual dispositions of men and animals, by the configuration of the brain and head, vol. 1.* Marsh, Capen & Lyon. [LMH]

Gandhi, S. P., Heeger, D. J. & Boynton, G. M. (1999) Spatial attention affects brain activity in human primary visual cortex. *Proceedings of the National Academy of Sciences USA* 96(6):3314–19. [DWV]

Ganong, W. F. (1980) Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance* 6:110–25. [ACu]

Gantman, A. P. & Van Bavel, J. J. (2014) The moral pop-out effect: Enhanced perceptual awareness of morally relevant stimuli. *Cognition* 132:22–29. [APG, arCF]

Gantman, A. P. & Van Bavel, J. J. (2015) Moral perception. *Trends in Cognitive Sciences* 19:631–33. [rCF, APG]

Gao, T. & Scholl, B. J. (2011) Chasing vs. stalking: Interrupting the perception of animacy. *Journal of Experimental Psychology: Human Perception and Performance* 37:669–84. [rCF]

Gao, T., McCarthy, G. & Scholl, B. J. (2010) The wolfpack effect: Perception of animacy irresistibly influences interactive behavior. *Psychological Science* 21:1845–53. [arCF]

Gao, T., Newman, G. E. & Scholl, B. J. (2009) The psychophysics of chasing: A case study in the perception of animacy. *Cognitive Psychology* 59(2):154–79. Available at: http://doi.org/10.1016/j.cogpsych.2009.03.001. [GL]

Garn, S. M., Leonard, W. R. & Hawthorne, V. M. (1986) Three limitations of the body mass index. *The American Journal of Clinical Nutrition* 44:996–97. [rCF]

Gau, R. & Noppeney, U. (2015) How prior expectations shape multisensory perception. *NeuroImage* 124(Pt A):876–86. doi:10.1016/j.neuroimage.2015.09.045. [BdH]

Gazzaley, A. & Nobre, A. C. (2012) Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Sciences* 16(2):129–35. [ACl]

Gentner, D. (2010) Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science* 34:752–75. [GSH]

Gerbino, W., Manzini, M., Rigutti, S. & Fantoni, C. (2014) Action molds the perception of facial expressions. *Perception, ECVP Abstract Supplement* 43:173. [WG]

Geuss, M. N., Stefanucci, J. K., de Benedictis-Kessner, J. & Stevens, N. R. (2010) A balancing act: Physical balance, through arousal, influences size perception. *Attention, Perception and Psychophysics* 72:1890–902. [aCF]

Ghazanfar, A. A. & Schroeder, C. E. (2006) Is neocortex essentially multisensory? *Trends in Cognitive Sciences* 10(6):278–85. Available at: http://dx.doi.org/10.1016/j.tics.2006.04.008) [TE]

Gibson, J. J. (1966) *The senses considered as perceptual systems*. Houghton Mifflin. [PT]

Gibson, J. J. (1979) *The ecological approach to perception and action*. Houghton Mifflin. [GLC, aCF, PT, JKW]

Gibson, J. J. (1979/1986) *The ecological approach to visual perception*. Erlbaum. (Original work published in 1979 by Houghton Mifflin.) [RC-B]

Gigerenzer, G. (2004) Mindless statistics. *Journal of Socio-Economics* 33:587–606. [rCF]

Gilbert, C. D. & Li, W. (2013) Top-down influences on visual processing. *Nature Reviews Neuroscience* 14(5):350–63. [DMB, APG, LMH, MR, aCF]

Gilbert, C. D. & Sigman, M. (2007) Brain states: Top-down influences in sensory processing. *Neuron* 54(5):677–96. [LMH]

Gilbert, C. D. & Wu, L. (2013) Top-down influences of visual processing. *Nature Reviews of Neuroscience* 14:350–63. [MG]

Gilchrist, A. & Jacobsen, A. (1984) Perception of lightness and illumination in a world of one reflectance. *Perception* 13:5–19. [aCF]

Gitelman, D. R., Penny, W. D., Ashburner, J. & Friston, K. J. (2003) Modeling regional and psychophysiologic interactions in fMRI: The importance of hemodynamic deconvolution. *NeuroImage* 19(1):200–207. [MD]

Gobell, J. & Carrasco, M. (2005) Attention alters the appearance of spatial frequency and gap size. *Psychological Science* 16:644–51. [aCF]

Goldstone, R. L. (1995) Effects of categorization on color perception. *Psychological Science* 6:298–304. [aCF]

Goldstone, R. L. (2000) Unitization during category learning. *Journal of Experimental Psychology: Human Perception and Performance* 26:86–112. [RG]

Goldstone, R. L. & Barsalou, L. W. (1998) Reuniting perception and conception. *Cognition* 65:231–62. [aCF]

Goldstone, R. L., de Leeuw, J. R. & Landy, D. H. (2015) Fitting perception in and to cognition. *Cognition* 135:24–29. [RG, aCF]

Goldstone, R. L. & Hendrickson, A. T. (2010) Categorical perception. *Wiley Interdisciplinary Reviews: Cognitive Science* 1:69–78. [WG]

Goldstone, R. L. & Steyvers, M. (2001) The sensitization and differentiation of dimensions during category learning. *Journal of Experimental Psychology: General* 130:116–39. [RG]

Goodale, M. A. & Milner, A. D. (1992) Separate visual pathways for perception and action. *Trends in Neurosciences* 15(1):20–25. [CO]

Goodwin, G. P. & Johnson-Laird, P. N. (2005) Reasoning about relations. *Psychological Review* 112:468–93. [GSH]

Gottlieb, J. (2012) Attention, learning, and the value of information. *Neuron* 76 (2):281–95. [LMH]

Graham, K. S., Barense, M. D. & Lee, A. C. H. (2010) Going beyond LTM in the MTL: A synthesis of neuropsychological and neuroimaging findings on the role of the medial temporal lobe in memory and perception. *Neuropsychologia* 48 (4):831–53. doi:10.1016/j.neuropsychologia.2010.01.001. [LLE]

Graham, R. & LaBar, K. S. (2012) Neurocognitive mechanisms of gaze-expression interactions in face processing and social attention. *Neuropsychologia* 50 (5):553–66. [CO]

Gray, R. (2002) Behavior of college baseball players in a virtual batting task. *Journal of Experimental Psychology: Human Perception and Performance* 28:1131–48. [RC-B]

Gray, R. (2010) Expert baseball batters have greater sensitivity in making swing decisions. *Research Quarterly for Exercise and Sport* 81:373–78. [RC-B]

Gray, R. (2013) Being selective at the plate: Processing dependence between perceptual variables relates to hitting goals and performance. *Journal of Experimental Psychology: Human Perception and Performance* 39:1124–42. [aCF]

Gray, R. & Cañal-Bruland, R. (2015) Attentional focus, perceived target size, and movement kinematics under performance pressure. *Psychonomic Bulletin and Review* 22:1692–1700. [aCF]

Gray, R., Navia, J. A. & Allsop, J. (2014) Action-specific effects in aviation: What determines judged runway size? *Perception* 43:145–54. [aCF]

Graziano, M. S. A. & Botvinick, M. (2002) How the brain represents the body: Insights from neurophysiology and psychology. In: *Common mechanisms in perception and action: Attention and performance XIX*, ed. W. Prinz & B. Hommel, pp. 136–57. Oxford University Press. [PT]

Gregory, R. L. (1980) Perceptions as hypotheses. *Philosophical Transactions of the Royal Society B: Biological Sciences* 290(1038):181–97. [DWV, aCF]

Gur, M. (2015) Space reconstruction by primary visual cortex activity patterns: A parallel, non-computational mechanism of object representation. *Trends in Neuroscience* 38:207–15. [MG]

Gur, M., Kagan, I. & Snodderly, M. (2005) Orientation and direction selectivity of single cells in V1 of alert monkeys: Functional relationships and laminar distributions. *Cerebral Cortex* 15:1207–21. [MG]

Gur, M. & Snodderly, D. M. (2008) Physiological differences between neurons in layer 2 and layer 3 of primary visual cortex (V1) of alert macaque monkeys. *Journal of Physiology (London)* 586:2293–306. [MG]

Haddad, R., Khan, R., Takahashi, Y. K., Mori, K., Harel, D. & Sobel, N. (2008) A metric for odorant comparison. *Nature Methods* 5(5):425–29. doi:10.1038/nmeth.1197. [AK]

Halford, G. S., Andrews, G., Phillips, S. & Wilson, W. H. (2013) The role of working memory in the subsymbolic–symbolic transition. *Current Directions in Psychological Science* 22:210–16. [GSH]

Halford, G. S., Baker, R., McCredden, J. E. & Bain, J. D. (2005) How many variables can humans process? *Psychological Science* 16:70–76. [GSH]

Halford, G. S., Wilson, W. H., Andrews, G. & Phillips, S. (2014) *Categorizing cognition: Toward conceptual coherence in the foundations of psychology*. MIT Press. [GSH]

Halford, G. S., Wilson, W. H. & Phillips, S. (2010) Relational knowledge: The foundation of higher cognition. *Trends in Cognitive Sciences* 14:497–505. [GSH]

Hansen, T., Giesel, M. & Gegenfurtner, K. R. (2008) Chromatic discrimination of natural objects. *Journal of Vision* 8(2):1–19. [aCF]

Hansen, T., Olkkonen, M., Walter, S. & Gegenfurtner, K. R. (2006) Memory modulates color appearance. *Nature Neuroscience* 9(11):1367–68. Available at: http://doi.org/10.1038/nn1794. [GL, CW, arCF]

Harber, K. D., Yeung, D. & Iacovelli, A. (2011) Psychosocial resources, threat, and the perception of distance and height: Support for the resources and perception model. *Emotion* 11:1080–90. [aCF]

Hart, W., Tullett, A. M., Shreves, W. B. & Fetterman, Z. (2015) Fueling doubt and openness: Experiencing the unconscious, constructed nature of perception induces uncertainty and openness to change. *Cognition* 137:1–8. [LMH]

Hartmann, M. & Fischer, M. H. (2014) Pupillometry: The eyes shed fresh light on the mind. *Current Biology* 24(7):R281–82. Available at: http://doi.org/10.1016/j.cub.2014.02.028. [WX]

Hay, J. & Drager, K. (2010) Stuffed toys and speech perception. *Linguistics* 48:865–92. [ACu]

Hay, J., Nolan, A. & Drager, K. (2006) From *fush* to *feesh*: Exemplar priming in speech perception. *The Linguistic Review* 23:351–79. [ACu]

Helmholtz, H. (1866/1925) *Handbuch der physiologischen optik. Leipzig: L. Voss. Translated as Treatise on physiological optics, vol. 3 (J. Southall, trans.)*. Optical Society of America. (Original work published in 1925). [aCF]

Hill, W. E. (1915) My wife and my mother-in-law. *Puck* Nov. 6, p. 11. [DWV]

Hillyard, S. A., Vogel, E. K. & Luck, S. J. (1998) Sensory gain control (amplification) as a mechanism of selective attention: Electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 353(1373):1257–70. [WX]

Hirstein, W. & Ramachandran, V. S. (1997) Capgras syndrome: A novel probe for understanding the neural representation of the identity and familiarity of persons. *Proceedings of the Royal Society B: Biological Sciences* 264 (1380):437–44. [RMR]

Hohwy, J. (2013) *The predictive mind*. Oxford University Press. [ACl]

Howard, I. P. & Rogers, B. J. (2002) *Seeing in depth, vol. 2: Depth perception*. University of Toronto Press. [aCF]

Howard, L. R., Kumaran, D., Ólafsdóttir, H. F. & Spiers, H. J. (2011) Double dissociation between hippocampal and parahippocampal responses to object-background context and scene novelty. *The Journal of Neuroscience* 31 (14):5253–61. doi:10.1523/JNEUROSCI.6055-10.2011. [LLE]

Hsieh, P.-J., Vul, E. & Kanwisher, N. (2010) Recognition alters the spatial pattern of FMRI activation in early retinotopic cortex. *Journal of Neurophysiology* 103:1501–507. Available at: http://doi.org/10.1152/jn.00812.2009. [DMB]

Hung, C. P., Kreiman, G., Poggio, T. & DiCarlo, J. J. (2005) Fast readout of object identity from macaque inferior temporal cortex. *Science* 310(5749):863–66. Available at: http://doi.org/10.1126/science.1117593. [DMB]

Hupé, J. M., James, A. C., Payne, B. R., Lomber, S. G., Girard, P. & Bullier, J. (1998) Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature* 394(6695):784–87. [DWV]

Huys, R., Cañal-Bruland, R., Hagemann, N., Beek, P. J., Smeeton, N. J. & Williams, A. M. (2009) Global information pickup underpins anticipation skill of tennis shot direction. *Journal of Motor Behavior* 41:158–70. [RC-B]

Ito, M., Tamura, H., Fujita, I. & Tanaka, K. (1995) Size and position invariance of neuronal responses in monkey inferotemporal cortex. *Journal of Neurophysiology* 73:218–26. [MG]

Jacobs, D. M. & Michaels, C. F. (2007) Direct learning. *Ecological Psychology* 19:321–49. [RC-B]

James, W. (1890) *The principles of psychology*. H. Holt. [GLC]

Jastrow, J. (1899) The mind's eye. *Popular Science Monthly* 54:299–312. [DWV]

Jehee, J. F. M., Ling, S., Swisher, J. D., van Bergen, R. S. & Tong, F. (2012) Perceptual learning selectively refines orientation representations in early visual cortex. *The Journal of Neuroscience* 32:16747–53. [RG]

Jordan, J. S. (2013) The wild ways of conscious will: What we do, how we do it, and why it has meaning. *Frontiers in Psychology* 4:574. [DWV]

Jussim, L., Crawford, J. T., Anglin, S. M., Stevens, S. T. & Duarte, J. L. (2016) Interpretations and methods: Towards a more effectively self-correcting social psychology. *Journal of Experimental Social Psychology* 66:116–33. [rCF]

Kanizsa, G. (1985) Seeing and thinking. *Acta Psychologica* 59:23–33. [arCF]

Kanizsa, G. & Gerbino, W. (1982) Amodal completion: Seeing or thinking? In: *Organization and representation in perception*, ed. J. Beck, pp. 167–90. Erlbaum. [rCF]

Kapadia, M. K., Ito, M., Gilbert, C. D. & Westheimer, G. (1995) Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. *Neuron* 15(4):843–56. [DWV]

Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R. & Ungerleider, L. G. (1999) Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron* 22(4):751–61. [DWV]

Kastner, S. & Ungerleider, L. G. (2001) The neural basis of biased competition in human visual cortex. *Neuropsychologia* 39(12):1263–76. [DWV]

Kennedy, B. L. & Most, S. B. (2012) Perceptual, not memorial, disruption underlies emotion-induced blindness. *Emotion* 12(2):199–202. [SBM]

Kennedy, B. L., Rawding, J., Most, S. B. & Hoffman, J. E. (2014) Emotion-induced blindness reflects competition for early and late processing stages: An ERP study. *Cognitive, Affective, and Behavioral Neuroscience* 14:1485–98. [SBM]

Keysers, C. & Perrett, D. I. (2002) Visual masking and RSVP reveal neural competition. *Trends in Cognitive Sciences* 6:120–25. [SBM]

Khemlani, S. & Johnson-Laird, P. N. (2012) Theories of the syllogism: A meta-analysis. *Psychological Bulletin* 138(3):427–57. [GSH]

Kiefer, M. & Barsalou, L. W. (2013) Grounding the human conceptual system in perception, action, and internal states. In: *Action science: Foundations of an emerging discipline*, ed. W. Prinz, M. Beisert & A. Herwig, pp. 381–407. MIT Press. [WG]

Kihara, K. & Takeda, Y. (2010) Time course of the integration of spatial frequency-based information in natural scenes. *Vision Research* 50:2158–62. [AR]

Kim, A. & Lai, V. (2012) Rapid interactions between lexical semantic and word form analysis during word recognition in context: Evidence from ERPs. *Journal of Cognitive Neuroscience* 24(5):1104–12. doi:10.1162/jocn_a_00148. [MR]

Kirsch, W. (2015) Impact of action planning on spatial perception: Attention matters. *Acta Psychologica* 156:22–31. [aCF]

Kirsch, W. & Kunde, W. (2013a) Moving further moves things further away in visual perception: Position-based movement planning affects distance judgments. *Experimental Brain Research* 226:431–40. [aCF]

Kirsch, W. & Kunde, W. (2013b) Visual near space is scaled to parameters of current action plans. *Journal of Experimental Psychology: Human Perception and Performance* 39:1313–25. [aCF]

Klein, B. P., Harvey, B. M. & Dumoulin, S. O. (2014) Attraction of position preference by spatial attention throughout human visual cortex. *Neuron* 84(1):227–37. doi:10.1016/j.neuron.2014.08.047. [BdH]

Klein, G. S., Schlesinger, H. J. & Meister, D. E. (1951) The effect of personal values on perception: An experimental critique. *Psychological Review* 58:96–12. [aCF]

Klein, I., Dubois, J., Mangin, J.-F., Kherif, F., Flandin, G., Poline, J.-B. & Le Bihan, D. (2004) Retinotopic organization of visual mental images as revealed by functional magnetic resonance imaging. *Cognitive Brain Research* 22(1):26–31. [TE]

Klemfuss, N., Prinzmetal, W. & Ivry, R. B. (2012) How does language change perception: A cautionary note. *Frontiers in Psychology* 3:Article 78. Available at: http://doi.org/10.3389/fpsyg.2012.00078. [GL, arCF]

Köhler, W. (1929) *Gestalt psychology*. Liveright. [WG]

Köhler, W. (1938) *The place of value in a world of facts*. Liveright. [WG]

Koivisto, M. & Revonsuo, A. (2007) How meaning shapes seeing. *Psychological Science* 18:845–49. [SBM]

Kok, P., Brouwer, G., van Gerven, M. & de Lange, F. (2013) Prior expectations bias sensory representations in visual cortex. *Journal of Neuroscience* 33:16275–84. [RO]

Kok, P., Jehee, J. F. M. & de Lange, F. P. (2012) Less is more: Expectation sharpens representations in the primary visual cortex. *Neuron* 75(2):265–70. Available at: http://doi.org/10.1016/j.neuron.2012.04.034. [DMB]

Kominsky, J. & Scholl, B. J. (2016) Retinotopic adaptation reveals multiple distinct categories of causal perception. *Journal of Vision* 16(12):333. [rCF]

Kosslyn, S. M. (1994) *Image and brain*. MIT Press. [AR]

Kosslyn, S. M. (2005) Mental images and the brain. *Cognitive Neuropsychology* 22:333–47. [aCF]

Kosslyn, S. M., Behrmann, M. & Jeannerod, M. (1995) The cognitive neuroscience of mental imagery. *Neuropsychologia* 33:1335–44. [PDLH]

Kosslyn, S. M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J., Ganis, G. & Alpert, N. (1999) The role of area 17 in visual imagery: Convergent evidence from PET and rTMS. *Science* 284(5411):167–70. [TE]

Kosslyn, S. M., Thompson, W. L. & Ganis, G. (2006) *The case for mental imagery*. Oxford University Press. [NB]

Krauskopf, J. & Gegenfurtner, K. (1992) Color discrimination and adaptation. *Vision Research* 32:2165–75. [aCF]

Kriegeskorte, N. & Kievit, R. A. (2013) Representational geometry: Integrating cognition, computation, and the brain. *Trends in Cognitive Sciences* 17:401–12. [VM]

Krpan, D. & Schnall, S. (2014) Too close for comfort: Stimulus valence moderates the influence of motivational orientation on distance perception. *Journal of Personality and Social Psychology* 107:978–93. [aCF]

Kveraga, K., Boshyan, J. & Bar, M. (2007a) Magnocellular projections as the trigger of top-down facilitation in recognition. *The Journal of Neuroscience* 27(48):13232–40. doi:10.1523/jneurosci.3481-07.2007. [CO]

Kveraga, K., Ghuman, A. S. & Bar, M. (2007b) Top-down predictions in the cognitive brain. *Brain and Cognition* 65(2):145–68. [DWV]

Kveraga, K., Ghuman, A. S., Kassam, K. S., Aminoff, E. A., Hämäläinen, M. S., Chaumon, M. & Bar, M. (2011) Early onset of neural synchronization in the contextual associations network. *Proceedings of the National Academy of Sciences USA* 108(8):3389–94. [CO]

Laeng, B. & Endestad, T. (2012) Bright illusions reduce the eye's pupil. *Proceedings of the National Academy of Sciences of the USA* 109(6):2162–67. Available at: http://doi.org/10.1073/pnas.1118298109. [WX]

Laeng, B., Sirois, S. & Gredebäck, G. (2012) Pupillometry: A window to the preconscious? *Perspectives on Psychological Science* 7(1):18–27. Available at: http://doi.org/10.1177/1745691611427305. [WX]

Laeng, B. & Sulutvedt, U. (2014) The eye pupil adjusts to imaginary light. *Psychological Science* 25(1):188–97. Available at: http://doi.org/10.1177/0956797613503556. [WX]

Landau, A. N., Aziz-Zadeh, L. & Ivry, R. B. (2010) The influence of language on perception: Listening to sentences about faces affects the perception of faces. *The Journal of Neuroscience* 30(45):15254–61. Available at: http://doi.org/10.1523/JNEUROSCI.2046-10.2010. [GL, aCF]

Landis, D., Jones, J. M. & Reiter, J. (1966) Two experiments on perceived size of coins. *Perceptual and Motor Skills* 23:719–29. [aCF]

Lavie, N. (2005) Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences* 9(2):75–82. doi:10.1016/j.tics.2004.12.004. [BdH]

Lazarus, R. S., Yousem, H. & Arenberg, D. (1953) Hunger and perception. *Journal of Personality* 21:312–28. [aCF]

Le Bihan, D., Turner, R., Zeffiro, T. A., Cuénod, C. A., Jezzard, P. & Bonnerot, V. (1993) Activation of human primary visual cortex during visual recall: A magnetic resonance imaging study. *Proceedings of the National Academy of Sciences USA* 90:11802–805. [aCF]

Lee, D. H., Mirza, R., Flanagan, J. G. & Anderson, A. K. (2014) Optical origins of opposing facial expression actions. *Psychological Science* 25(3):745–52. Available at: http://doi.org/10.1177/0956797613514451. [WX, aCF]

Lee, S.-H., Blake, R. & Heeger, D. J. (2005) Traveling waves of activity in primary visual cortex during binocular rivalry. *Nature Neuroscience* 8(1):22–23. Available at: http://doi.org/10.1038/nn1365. [DMB]

Léger, L. & Chauvet, E. (2015) When canary primes yellow: Effects of semantic memory on overt attention. *Psychonomic Bulletin and Review* 22:200–205. [aCF]

Leopold, D. A. & Logothetis, N. K. (1996) Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature* 379:549–53. Available at: http://doi.org/10.1038/379549a0. [DMB]

Lessard, D. A., Linkenauger, S. A. & Proffitt, D. R. (2009) Look before you leap: Jumping ability affects distance perception. *Perception* 38:1863–66. [aCF]

Levin, D. T. & Banaji, M. R. (2006) Distortions in the perceived lightness of faces: The role of race categories. *Journal of Experimental Psychology: General* 135:501–12. [arCF, DTL]

Levin, D. T., Takarae, Y., Miner, A. G. & Keil, F. (2001) Efficient visual search by category: Specifying the features that mark the difference between artifacts and animals in preattentive vision. *Perception and Psychophysics* 63:676–97. [arCF]

Lewis, K., Borst, G. & Kosslyn, S. M. (2011) Integrating images and percepts: New evidence for depictive representation. *Psychological Research* 75(4):259–71. [NB]

Li, J.-L. & Yeh, S.-L. (2003) Do "Chinese and American see opposite apparent motions in a Chinese character"? Tse and Cavanagh (2000) replicated and revised. *Visual Cognition* 10:537–47. [aCF]

Li, W., Piëch, V. & Gilbert, C. D. (2006) Contour saliency in primary visual cortex. *Neuron* 50(6):951–62. [DMB]

Li, Z. & Durgin, F. H. (2011) Design, data and theory regarding a digital hand inclinometer: A portable device for studying slant perception. *Behavior Research Methods* 43:363–71. [FHD]

Lim, S. L., Padmala, S. & Pessoa, L. (2009) Segregating the significant from the mundane on a moment-to-moment basis via direct and indirect amygdala contributions. *Proceedings of the National Academy of Sciences USA* 106:16841–46. [APG]

Lindquist, K. A. & Barrett, L. F. (2012) A functional architecture of the human brain: Emerging insights from the science of emotion. *Trends in Cognitive Sciences* 16(11):533–40. [LMH]

Liu, T., Abrams, J. & Carrasco, M. (2009) Voluntary attention enhances contrast appearance. *Psychological Science* 20(3):354–62. [BdH]

Lo Sciuto, L. A. & Hartley, E. L. (1963) Religious affiliation and open-mindedness in binocular resolution. *Perceptual and Motor Skills* 17:427–30. [aCF]

Logothetis, N. K, Pauls, J. & Poggio, T. (1995) Shape representation in the inferior cortex of monkeys. *Current Biology* 5:552–63. [RG]

Logothetis, N. K. & Schall, J. D. (1989) Neuronal correlates of subjective visual perception. *Science* 245(4919):761–63. Available at: http://doi.org/10.1126/science.2772635. [DMB]

Long, G. M. & Toppino, T. C. (2004) Enduring interest in perceptual ambiguity: Alternating views of reversible figures. *Psychological Bulletin* 130:748–68. [arCF]

Lu, Z., Guo, B., Boguslavsky, A., Cappiello, M., Zhang, W. & Meng, M. (2015) Distinct effects of contrast and color on subjective rating of fearfulness. *Frontiers in Psychology* 6:1–9. Available at: http://doi.org/10.3389/fpsyg.2015.01521. [WX]

Luck, S. J. (1995) Multiple mechanisms of visual-spatial attention: Recent evidence from human electrophysiology. *Behavioral and Brain Research* 71:113–23. [AR]

Lupyan, G. (2012) Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in Psychology* 3:Article 54. [arCF]

Lupyan, G. (2015a) Cognitive penetrability of perception in the age of prediction: Predictive systems are penetrable systems. *Review of Philosophy and Psychology* 6:547–69. [NB, GL]

Lupyan, G. (2015b) Object knowledge changes visual appearance: Semantic effects on color afterimages. *Acta Psychologica* 161:117–30. [GL]

Lupyan, G. & Clark, A. (2015) Words and the world: Predictive coding and the language-perception-cognition interface. *Current Directions in Psychological Science* 24(4):279–84. [GL]

Lupyan, G. & Spivey, M. J. (2008) Perceptual processing is facilitated by ascribing meaning to novel stimuli. *Current Biology* 18:R410–12. [arCF, GL]

Lupyan, G. & Spivey, M. J. (2010) Making the invisible visible: Verbal but not visual cues enhance visual detection. *PLoS ONE* 5(7):e11452. [DWV]

Lupyan, G., Thompson-Schill, S. L. & Swingley, D. (2010) Conceptual penetration of visual processing. *Psychological Science* 21(5):682–91. [GL, arCF]

Lupyan, G. & Ward, E. J. (2013) Language can boost otherwise unseen objects into visual awareness. *Proceedings of the National Academy of Sciences USA* 110 (35):14196–201. Available at: http://doi.org/10.1073/pnas.1303312110. [GL, arCF]

Lyons, J. (2011) Circularity, reliability, and the cognitive penetrability of perception. *Philosophical Issues* 21:289–311. [WG]

Machery, E. (2015) Cognitive penetrability: A no-progress report. In: *The cognitive penetrability of perception: New philosophical perspectives*, ed. J. Zeimbekis & A. Raftopoulos. Oxford University Press. [aCF]

Mack, A. & Rock, I. (1998) *Inattentional blindness*. MIT Press. [SBM]

MacLeod, C., Mathews, A. & Tata, P. (1986) Attentional bias in emotional disorders. *Journal of Abnormal Psychology* 95(1):15. [SBM]

Macpherson, F. (2012) Cognitive penetration of colour experience: Rethinking the issue in light of an indirect mechanism. *Philosophy and Phenomenological Research* 84(1):24–62. [NB, arCF]

Macpherson, F. (2015) Cognitive penetration and nonconceptual content. In: *The cognitive penetrability of perception: New philosophical perspectives*, ed. J. Zeimbekis & A. Raftopoulos. pp. 331–59. Oxford University Press. [CO]

Makino, H. & Komiyama, T. (2015) Learning enhances the relative impact of top-down processing in the visual cortex. *Nature Neuroscience* 18(8):1116–22. [LMH]

Maloney, L. T. & Mamassian, P. (2009) Bayesian decision theory as a model of human visual perception: Testing Bayesian transfer. *Visual Neuroscience* 26 (1):147–55. [CW]

Maner, J. K., Kenrick, D. T., Becker, D. V., Robertson, T. E., Hofer, B., Neuberg, S. L., Delton, A.W., Butner, J. & Schaller, M. (2005) Functional projection: How fundamental social motives can bias interpersonal perception. *Journal of Personality and Social Psychology* 88:63–78. [aCF]

Mangun, G. R. & Hilyard, S. A. (1995) Electrophysiological studies of visual selective attention in humans. In: *The neurobiology of higher cognitive function*, ed. M. D. Rugg & M. G. Coles, pp. 40–85. Oxford University Press. [AR]

Mann, D. L., Abernethy, B. & Farrow, D. (2010) Action specificity increases anticipatory performance and the expert advantage in natural interceptive tasks. *Acta Psychologica* 135:17–23. [RC-B]

Mann, D. T. Y., Williams, A. M., Ward, P. & Janelle, C. M. (2007) Perceptual-cognitive expertise in sport: A meta-analysis. *Journal of Sport and Exercise Psychology* 29:457–78. [RC-B]

Marder, E. (2012) Neuromodulation of neuronal circuits: Back to the future. *Neuron* 76(1):1–11. [LMH]

Markant, J. & Amso, D. (2013) Selective memories: Infants' encoding is enhanced in selection via suppression. *Developmental Science* 16(6):926–40. [DWV]

Markant, J. & Amso, D. (2016) The development of selective attention orienting is an agent of change in learning and memory efficacy. *Infancy* 21(2): 154–76. doi: 10.1111/infa.12100. [DWV]

Markant, J., Oakes, L. M. & Amso, D. (2015a) Visual selective attention biases contribute to the other-race effect among 9-month-old infants. *Developmental Psychobiology* 58(3):355–65. doi:10.1002/dev.21375. [DWV]

Markant, J., Worden, M. S. & Amso, D. (2015b) Not all attention orienting is created equal: Recognition memory is enhanced when attention orienting involves distractor suppression. *Neurobiology of Learning and Memory* 120:28–40. [DWV]

Markov, N. T. & Kennedy, H. (2013) The importance of being hierarchical. *Current Opinion in Neurobiology* 23:187–94. [VM]

Martín, A., Chambeaud, J. G. & Barraza, J. F. (2015) The effect of object familiarity on the perception of motion. *Journal of Experimental Psychology. Human Perception and Performance* 41(2):283–88. Available at: http://doi.org/10.1037/xhp0000027. [GL]

Masson, M. E. & Borowsky, R. (1998) More than meets the eye: Context effects in word identification. *Memory and Cognition* 26:1245–69. [aCF]

Masters, R., Poolton, J. & van der Kamp, J. (2010) Regard and perceptions of size in soccer: Better is bigger. *Perception* 39:1290–95. [aCF]

Mathôt, S., van der Linden, L., Grainger, J. & Vitu, F. (2015) The pupillary light response reflects eye-movement preparation. *Journal of Experimental Psychology: Human Perception and Performance* 41(1):28–35. Available at: http://doi.org/10.1037/a0038653. [WX]

McCurdy, H. G. (1956) Coin perception studies and the concept of schemata. *Psychological Review* 63:160–68. [aCF]

McGann, J. P. (2015) Associative l earning and sensory neuroplasticity: How does it happen and what is it good for? *Learning & Memory* 22:567–76. [VM]

McGurk, H. & MacDonald, J. (1976) Hearing lips and seeing voices. *Nature* 264:746–48. [rCF, BdH, ACu]

McQueen, J. M. (1991) The influence of the lexicon on phonetic categorization: Stimulus quality in word-final ambiguity. *Journal of Experimental Psychology: Human Perception and Performance* 17:433–43. [ACu]

Meier, B. P., Robinson, M. D., Crawford, L. E. & Ahlvers, W. J. (2007) When "light" and "dark" thoughts become light and dark responses: Affect biases brightness judgments. *Emotion* 7(2):366–76. Available at: http://doi.org/10.1037/1528-3542.7.2.366. [WX, PT, arCF]

Meijer, P. B. (1992) An experimental system for auditory image representations. *IEEE Transactions on Biomedical Engineering* 39(2):112–21. doi:10.1109/10.121642. [TE]

Meredith, M. A. & Stein, B. E. (1986) Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology* 56(3):640–62. [TE]

Meteyard, L., Bahrami, B. & Vigliocco, G. (2007) Motion detection and motion verbs language affects low-level visual perception. *Psychological Science* 18:1007–13. [aCF]

Metzger, W. (1941) *Psychologie*. Steinkopff. [WG]

Meyer, D. E. & Schvaneveldt, R. W. (1971) Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology* 90:227–34. [aCF]

Milner, A. D. & Goodale, M. A. (2008) Two visual systems reviewed. *Neuropsychologia* 54:774–85. [RC-B]

Miskovic, V. & Keil, A. (2012) Acquired fears reflected in cortical sensory processing: A review of electrophysiological studies of human classical conditioning. *Psychophysiology* 49:1230–41. [VM]

Mitterer, H., Horschig, J. M., Müsseler, J. & Majid, A. (2009) The influence of memory on perception: It's not what things look like, it's what you call them. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35:1557–62. [aCF]

Mole, C. (2015) Attention and cognitive penetration. In: *The cognitive penetrability of perception: New philosophical perspectives*, ed. J. Ziembekis & A. Raftopolous, pp. 218–38. Oxford University Press. [AR]

Montagna, B., Pestilli, F. & Carrasco, M. (2009) Attention trades off spatial acuity. *Vision Research* 49(7):735–45. [BdH]

Moore, C. & Cavanagh, P. (1998) Recovery of 3D volume from 2-tone images of novel objects. *Cognition* 67(1–2):45–71. Available at: http://doi.org/10.1016/S0010-0277(98)00014-6. [GL]

Moore, T. & Armstrong, K. M. (2003) Selective gating of visual signals by microstimulation of frontal cortex. *Nature* 421(6921):370–73. Available at: http://doi.org/10.1038/nature01341. [DMB]

Morgan, M. J. (1977) *Molyneux's question: Vision, touch and the philosophy of perception*. Cambridge University Press. [TE]

Most, S. B. (2010) What's "inattentional" about inattentional blindness? *Consciousness and Cognition* 19:1102–104. [SBM]

Most, S. B. (2013) Setting sights higher: Category-level attentional set modulates sustained inattentional blindness. *Psychological Research* 77:139–46. [SBM, rCF]

Most, S. B., Simons, D. J., Scholl, B. J., Jimenez, R., Clifford, E. & Chabris, C. F. (2001) How not to be seen: The contribution of similarity and selective ignoring to sustained inattentional blindness. *Psychological Science* 12(1):9–17. [SBM, rCF]

Most, S. B., Chun, M. M., Widders, D. M. & Zald, D. H. (2005a) Attentional rubbernecking: Cognitive control and personality in emotion-induced blindness. *Psychonomic Bulletin and Review* 12:654–61. [SBM]

Most, S. B., Laurenceau, J.-P., Graber, E., Belcher, A. & Smith, C. V. (2010) Blind jealousy: Romantic insecurity increases emotion-induced failures of visual perception. *Emotion* 10:250–56. [SBM]

Most, S. B., Scholl, B. J., Clifford, E. R. & Simons, D. J. (2005b) What you see is what you set: Sustained inattentional blindness and the capture of awareness. *Psychological Review* 112:217–42. [SBM, arCF]

Most, S. B. & Wang, L. (2011) Dissociating spatial attention and awareness in emotion-induced blindness. *Psychological Science* 22:300–305. [rCF, SBM]

Motter, B. C. (1993) Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *Journal of Neurophysiology* 70(3):909–19. Available at: http://doi.org/0022-3077/93. [DMB, DWV]

Muckli, L. (2010) What are we missing here? Brain imaging evidence for higher cognitive functions in primary visual cortex V1. *International Journal of Imaging Systems and Technology* 20:131–39. [VM]

Muckli, L. & Petro, L. S. (2013) Network interactions: Non-geniculate input to V1. *Current Opinion in Neurobiology* 23:195–201. [VM]

Murphy, P. R., Vandekerckhove, J. & Nieuwenhuis, S. (2014) Pupil-linked arousal determines variability in perceptual decision making. *PLoS Computational Biology* 10(9):e1003854. Available at: http://doi.org/10.1371/journal.pcbi.1003854. [WX]

Naber, M. & Nakayama, K. (2013) Pupil responses to high-level image content. *Journal of Vision* 13(6):7. Available at: http://doi.org/10.1167/13.6.7. [WX]

Nanay, B. (2014) Cognitive penetration and the gallery of indiscernibles. *Frontiers in Psychology* 5:1527. [aCF]

Naselaris, T., Olman, C. A., Stansbury, D. E., Ugurbil, K. & Gallant, J. L. (2015) A voxel-wise encoding model for early visual areas decodes mental images of remember scenes. *NeuroImage* 105:215–28. [PDLH]

Navarra, J., Alsius, A., Soto-Faraco, S. & Spence, C. (2010) Assessing the role of attention in the audiovisual integration of speech. *Information Fusion* 11(1):4–11. doi:10.1016/j.inffus.2009.04.001. [BdH]

Nayar, K., Franchak, F., Adolph, K. & Kiorpes, L. (2015) From local to global processing: The development of illusory contour perception. *Journal of Experimental Child Psychology* 131:38–55. [GSH]

New, J. J. & German, T. C. (2015) Spiders at the cocktail party: An ancestral threat that surmounts inattentional blindness. *Evolution and Human Behavior* 36:165–73. [aCF]

Newsome, W. T., Britten, K. H. & Movshon, J. A. (1989) Neuronal correlates of a perceptual decision. *Nature* 341(6237):52–54. Available at: http://doi.org/10.1038/341052a0. [DMB]

Niedenthal, P. M., Mermillod, M., Maringer, M. & Hess, U. (2010) The Simulation of Smiles (SIMS) model: Embodied simulation and the meaning of facial expression. *Behavioral and Brain Sciences* 33:417–80. [WG]

Niedzielski, N. (1999) The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18:62–85. [ACu]

Nobre, A., Griffin, I. & Rao, A. (2008) Spatial attention can bias search in visual short-term memory. *Frontiers in Human Neuroscience* 1:1–9. [ACl]

Norris, D. (1995) Signal detection theory and modularity: On being sensitive to the power of bias models of semantic priming. *Journal of Experimental Psychology: Human Perception and Performance* 21:935–39. [aCF]

Norris, D., McQueen, J. M. & Cutler, A. (2000) Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences* 23:299–370. [rCF, ACu]

Norris, D., McQueen, J. M. & Cutler, A. (2016) Prediction, Bayesian inference and feedback in speech recognition. *Language, Cognition and Neuroscience* 31 (1):4–18. doi:10.1080/23273798.2015.1081703. [ACu]

O'Connor, D. H., Fukui, M. M., Pinsk, M. A. & Kastner, S. (2002) Attention modulates responses in the human lateral geniculate nucleus. *Nature Neuroscience* 5 (11):1203–209. Available at: http://doi.org/10.1038/nn957. [DMB]

Oberauer, K. (2009) Design for a working memory. In: *Psychology of learning and motivation: Advances in research and theory, vol. 51*, ed. B. H. Ross, pp. 45–100. Elsevier, Academic Press. [GSH]

Ogilvie, R. & Carruthers, P. (2016) Opening up vision: The case against encapsulation. *Review of Philosophy and Psychology* 7:1–22 [ePub ahead of print]. doi:10.1007/s13164-015-0294-8. [RO]

Olkkonen, M., Hansen, T. & Gegenfurtner, K. R. (2008) Color appearance of familiar objects: Effects of object shape, texture, and illumination changes. *Journal of Vision* 8(5):1–16. [CW, aCF]

Olshausen, B. A. & Field, D. J. (2005) How close are to understanding V1? *Neural Computation* 17:1665–99. [VM]

Orne, M. T. (1962) On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist* 17:776–83. [rCF]

Pan, Y., Lin, B., Zhao, Y. & Soto, D. (2014) Working memory biasing of visual perception without awareness. *Attention, Perception and Psychophysics* 76:2051–62. [aCF]

Panichello, M. F., Cheung, O. S. & Bar, M. (2012) Predictive feedback and conscious visual experience. *Frontiers in Psychology* 3(620):1–8. [CO]

Pascual-Leone, A. & Walsh, V. (2001) Fast backprojections from the motion to the primary visual area necessary for visual awareness. *Science* 292(5516):510–12. Available at: http://doi.org/10.1126/science.1057099. [DMB]

Pascucci, D. & Turatto, M. (2013) Immediate effect of internal reward on visual adaptation. *Psychological Science* 24:1317–22. [aCF]

Pearson, J., Rademaker, R. L. & Tong, F. (2011) Evaluating the mind's eye: The metacognition of visual imagery. *Psychological Science* 22:1535–42. [PDLH]

Pelekanos, V. & Moutoussis, K. (2011) The effect of language on visual contrast sensitivity. *Perception* 40(12):1402–12. Available at: http://doi.org/10.1068/p7010. [GL]

Penny, W. D., Stephan, K. E., Daunizeau, J., Rosa, M. J., Friston, K. J., Schofield, T. M. & Leff, A. P. (2010) Comparing families of dynamic causal models. *PLoS Computational Biology* 6(3):e1000709. [MD]

Pessoa, L. (2015) Précis on *The Cognitive-Emotional Brain*. *Behavioral and Brain Sciences* 38:e71. [LMH]

Peters, A. (2002) Examining neocortical circuits: Some background and facts. *Journal of Neurocytology* 31(3–5):183–93. [LMH]

Peterson, M. A. & Gibson, B. S. (1991) Directing spatial attention within an object: Altering the functional equivalence of shape descriptions. *Journal of Experimental Psychology: Human Perception and Performance* 17:170–82. [arCF]

Peterson, M. A. & Gibson, B. S. (1993) Shape recognition inputs to figure-ground organization in three-dimensional displays. *Cognitive Psychology* 25:383–29. [aCF]

Peterson, M. A. & Gibson, B. S. (1994) Object recognition contributions to figure-ground organization: Operations on outlines and subjective contours. *Perception and Psychophysics* 56:551–64. [aCF]

Pezzulo, G. (2012) An active inference view of cognitive control. *Frontiers in Psychology* 3:478. [LMH]

Philbeck, J. W. & Witt, J. K. (2015) Action-specific influences on perception and post-perceptual processes: Present controversies and future directions. *Psychological Bulletin* 141(6):1120–44. doi:10.1037/a0039738 [JKW]

Phillips, S. & Wilson, W. H. (2014) A category theory explanation for systematicity: Universal constructions. In: *The architecture of cognition: Rethinking Fodor and Pylyshyn's systematicity challenge*, ed. P. Calvo & J. Symons, pp. 227–49. MIT Press. [GSH]

Pilgramm, S., de Haas, B., Helm, F., Zentgraf, K., Stark, R., Munzert, J. & Krüger, B. (2016) Motor imagery of hand actions: Decoding the content of motor imagery from brain activity in frontal and parietal motor areas. *Human Brain Mapping* 37(1):81–93. doi:10.1002/hbm.23015. [BdH]

Pinker, S. (1997) *How the mind works*. Norton. [LMH]

Pinna, B. & Brelstaff, G. J. (2000) A new visual illusion of relative motion. *Vision Research* 40:2091–96. [rCF]

Pitt, M. A. & McQueen, J. M. (1998) Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language* 39:347–70. [ACu]

Pitts, S., Wilson, J. P. & Hugenberg, K. (2014) When one is ostracized, others loom: Social rejection makes other people appear closer. *Social Psychological and Personality Science* 5:550–57. [aCF]

Pratte, M. S. & Tong, F. (2014) Spatial specificity of working memory representations in the early visual cortex. *Journal of Vision* 14(3): 22. Available at: http://doi.org/10.1167/14.3.22. [GL]

Prinz, J. & Seidel, A. (2012) Alligator or squirrel: Musically induced fear reveals threat in ambiguous figures. *Perception* 41:1535–39. [aCF]

Proffitt, D. R. (2006) Embodied perception and the economy of action. *Perspectives on Psychological Science* 1:110–22. [GLC, aCF]

Proffitt, D. R. (2009) Affordances matter in geographical slant perception. *Psychonomic Bulletin and Review* 16:970–72. [GLC]

Proffitt, D. R. (2013) An embodied approach to perception: By what units are Visual perceptions scaled? *Perspectives on Psychological Science* 8:474–83. [GLC]

Proffitt, D. R., Bhalla, M., Gossweiler, R. & Midgett, J. (1995) Perceiving geographical slant. *Psychonomic Bulletin and Review* 2:409–28. [GLC]

Proffitt, D. R. & Linkenauger, S. A. (2013) Perception viewed as a phenotypic expression. In: *Action science: Foundations of an emerging discipline*, ed. W. Prinz, M. Beisert & A. Herwig, pp. 171–98. MIT Press. [GLC, aCF]

Proffitt, D. R., Stefanucci, J. K., Banton, T. & Epstein, W. (2003) The role of effort in perceiving distance. *Psychological Science* 14:106–12. [aCF]

Proulx, M. J. (2011) Consciousness: What, how, and why. *Science* 332(6033):1034–35. doi:10.1126/Science.1206704. [TE]

Proulx, M. J., Brown, D. J., Pasqualotto, A. & Meijer, P. (2014) Multisensory perceptual learning and sensory substitution. *Neuroscience and Biobehavoral Reviews* 41:16–25. doi:10.1016/j.neubiorev.2012.11.017. [TE]

Pylyshyn, Z. (1999) Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences* 22(3):341–65. [DWV, arCF]

Pylyshyn, Z. W. (2002) Mental imagery: In search of a theory. *Behavioral and Brain Sciences* 25:157–238. [rCF]

Quinn, P. C. & Bhatt, R. S. (2005) Good continuation affects discrimination of visual pattern information in young infants. *Perception and Psychophysics* 67:1171–76. [GSH]

Rabovsky, M., Sommer, W. & Abdel Rahman, R. (2011) Depth of conceptual knowledge modulates visual processes during word reading. *Journal of Cognitive Neuroscience* 24(4):990–1005. doi:10.1162/jocn_a_00117. [MR]

Radel, R. & Clément-Guillotin, C. (2012) Evidence of motivational influences in early visual perception: Hunger modulates conscious access. *Psychological Science* 23:232–34. [APG, aCF]

Raftopoulos, A. (2001a) Is perception informationally encapsulated? The issue of the theory-ladenness of perception. *Cognitive Science* 25:423–51. [aCF]

Raftopoulos, A. (2001b) Reentrant neural pathways and the theory-ladenness of perception. *Philosophy of Science* 68:S187–99. [aCF]

Raftopoulos, A. (2009) *Cognition and perception: How do psychology and neural science inform philosophy*? MIT Press. [AR]

Raftopoulos, A. (2011) Late vision: Its processes and epistemic status. *Frontiers in Psychology* 2(382):1–12. doi:10.3389/fpsyg.2011.00382. [AR]

Raftopoulos, A. (2015) The cognitive impenetrability of perception and theory-ladenness. *Journal of General Philosophy of Science* 46(1):87–103. [AR]

Ramachandran, V. (1988) Perception of shape from shading. *Nature* 331:163–66. [aCF]

Rao, R. P. & Ballard, D. H. (1999) Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience* 2(1):79–87. [LMH, DWV]

Recanzone, G. H., Schreiner, C. E. & Merzenich, M. M. (1993) Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *Journal of Neuroscience* 13:87–103. [RG]

Rees, G., Frith, C. D. & Lavie, N. (1997) Modulating irrelevant motion perception by varying attentional load in an unrelated task. *Science* 278(5343):1616–19. [BdH]

Remez, R. E., Pardo, J. S., Piorkowski, R. L. & Rubin, P. E. (2001) On the bistability of sine wave analogues of speech. *Psychological Science* 12(1):24–29. [DWV]

Remez, R. E., Rubin, P. E., Pisoni, D. B. & Carrell, T. D. (1981) Speech perception without traditional speech cues. *Science* 212:947–49. [rCF]

Riccio, M., Cole, S. & Balcetis, E. (2013) Seeing the expected, the desired, and the feared: Influences on perceptual interpretation and directed attention. *Social and Personality Psychology Compass* 7:401–14. [aCF]

Riener, C. R., Stefanucci, J. K., Proffitt, D. R. & Clore, G. L. (2011) An effect of mood on the perception of geographical slant. *Cognition and Emotion* 25:174–82. [aCF]

Riskind, J. H., Moore, R. & Bowley, L. (1995) The looming of spiders: The fearful perceptual distortion of movement and menace. *Behaviour Research and Therapy* 33:171–78. [aCF]

Rock, I. (1983) *The logic of perception*. MIT Press. [aCF]

Rockland, K. S. (1994) The organization of feedback connections from area V2 (18) to V1 (17). In: *Primary visual cortex in primates*, ed. A. Peters & K. S. Rockland, pp. 261–99. Plenum Press. [MG]

Roelfsema, P. R., Lamme, V. A. F. & Spekreijse, H. (1998) Object-based attention in the primary visual cortex of the macaque monkey. *Nature* 395:376–81. [AR]

Roelfsema, P. R. & Spekreijse, H. (2001) The representation of erroneously perceived stimuli in the primary visual cortex. *Neuron* 31(5):853–63. Available at: http://doi.org/10.1016/S0896-6273(01)00408-1. [DMB]

Roepstorff, A. & Frith, C. (2004) What's at the top in the top-down control of action? Script-sharing and "top-top" control of action in cognitive experiments. *Psychological Research* 68(2–3):189–98. [DWV]

Rolfs, M., Dambacher, M. & Cavanagh, P. (2013) Visual adaptation of the perception of causality. *Current Biology* 23(3):250–54. doi:10.1016/j.cub.2012.12.017. [MR, arCF]

Rolls, E. T. (2008) Top–down control of visual perception: Attention in natural vision. *Perception* 37:333–54. [aCF]

Romei, V., de Haas, B., Mok, R. M. & Driver, J. (2011) Auditory stimulus timing influences perceived duration of co-occurring visual stimuli. *Frontiers in Psychology* 2:215. doi:10.3389/fpsyg.2011.00215. [BdH]

Rosenthal, R. & Rubin, D. B. (1978) Interpersonal expectancy effects: The first 345 studies. *Behavioral and Brain Sciences* 1:377–86. [aCF]

Ross, L. & Ward, A. (1996) Naïve realism in everyday life: Implications for social conflict and misunderstanding. In: *Values and knowledge*, ed. E. S. Reed, E. Turiel & T. Brown, pp. 103–35. Erlbaum. [LMH]

Russell, B. (1914) On the nature of acquaintance. II. Neutral monism. *The Monist* 24 (2):161–87. [RMR]

Rust, N. C., Mante, V., Simoncelli, E. P. & Movshon, J. A. (2006) How MT cells analyze the motion of visual patterns. *Nature Neuroscience* 9:1421–31. [aCF]

Rutkowski, R. G. & Weinberger, N. M. (2005) Encoding of learned importance of sound by magnitude of representational area in primary auditory cortex. *Proceedings of the National Academy of Sciences USA* 102:13664–69. [VM]

Rutman, A. M., Clapp, W. C., Chadick, J. Z. & Gazzaley, A. (2010) Early top–down control of visual processing predicts working memory performance. *Journal of Cognitive Neuroscience* 22(6):1224–34. [DWV]

Sacks, O. (2006) A neurologist's notebook, "Stereo Sue," *The New Yorker*, 82 (18):64–73. [RG]

Salin, P.-A. & Bullier, J. (1995) Corticocortical connections in the visual system: Structure and function. *Physiological Reviews* 75(1):107–54. [CO]

Schnall, S., Zadra, J. R. & Proffitt, D. R. (2010) Direct evidence for the economy of action: Glucose and the perception of geographical slant. *Perception* 39:464–82. [GLC, FHD, arCF]

Scholl, B. J. (2001) Objects and attention: The state of the art. *Cognition* 80:1–46. [aCF]

Scholl, B. J. (2007) Object persistence in philosophy and psychology. *Mind and Language* 22:563–91. [rCF]

Scholl, B. J. & Gao, T. (2013) Perceiving animacy and intentionality: Visual processing or higher-level judgment? In: *Social perception: Detection and interpretation of animacy, agency, and intention*, ed. M. D. Rutherford & V. A. Kuhlmeier, pp. 197–230. MIT Press. [arCF]

Scholl, B. J. & Leslie, A. M. (1999) Modularity, development and "theory of mind." *Mind and Language* 14:131–53. [aCF]

Scholl, B. J. & Nakayama, K. (2002) Causal capture: Contextual effects on the perception of collision events. *Psychological Science* 13:493–98. [rCF]

Scholl, B. J. & Tremoulet, P. D. (2000) Perceptual causality and animacy. *Trends in Cognitive Sciences* 4:299–309. [aCF]

Schultz, G. & Melzack, R. (1991) The Charles Bonnet syndrome: "Phantom visual images." *Perception* 20:809–25. [PDLH]

Schultz, G., Needham, W., Taylor, R., Shindell, S. & Melzack, R. (1996) Properties of complex hallucinations associated with deficits in vision. *Perception* 25:715–26. [PDLH]

Schwartz, S., Vuilleumier, P., Hutton, C., Maravita, A., Dolan, R. J. & Driver, J. (2005) Attentional load and sensory competition in human vision: Modulation of fMRI responses by load at fixation during task-irrelevant stimulation in the peripheral visual field. *Cerebral Cortex* 15:770–86. [MD]

Schwarz, N. (2015) Metacognition. In: *APA handbook of personality and social psychology: Vol. 1, Attitudes and social cognition*, ed. E. Borgida & J. A. Bargh, pp. 203–29. American Psychological Association. [GLC]

Scocchia, L., Cicchini, G. M. & Triesch, J. (2013) What's "up"? Working memory contents can bias orientation processing. *Vision Research* 78:46–55. Available at: http://doi.org/10.1016/j.visres.2012.12.003. [GL]

Scolari, M., Byers, A. & Serences, J. T. (2012) Optimal deployment of attentional gain during fine discriminations. *Journal of Neuroscience* 32(22):7723–33. Available at: http://doi.org/10.1523/JNEUROSCI.5558-11.2012. [DMB]

Sekuler, R., Sekuler, A. B. & Lau, R. (1997) Sound alters visual motion perception. *Nature* 385:308. [aCF]

Selimbeyoglu, A. & Parvizi, J. (2010) Electrical stimulation of the human brain: Perceptual and behavioral phenomena reported in the old and new literature. *Frontiers in Human Neuroscience* 4(46):1–11. [CO]

Shaffer, D. M. & Flint, M. (2011) Escalating slant: Increasing physiological potential does not reduce slant overestimates. *Psychological Science* 22:209–11. [aCF]

Shaffer, D. M., McManama, E., Swank, C. & Durgin, F. H. (2013) Sugar and space? Not the case: Effects of low blood glucose on slant estimation are mediated by beliefs. *i-Perception* 4:147–55. [FHD, aCF]

Shaffer, D. M., McManama, E., Swank, C., Williams, M. & Durgin, F. H. (2014) Anchoring in action: Manual estimates of slant are powerfully biased toward initial hand orientation and are correlated with verbal report. *Journal of Experimental Psychology: Human Perception and Performance* 40:1203–12. [FHD]

Shams, L., Kamitani, Y. & Shimojo, S. (2000) Illusions: What you see is what you hear. *Nature* 408:788. [BdH, arCF]

Shams, L., Kamitani, Y. & Shimojo, S. (2002) Visual illusion induced by sound. *Cognitive Brain Research* 14:147–52. [aCF]

Shiffrin, R. M. & Lightfoot, N. (1997) Perceptual learning of alphanumeric-like characters. *The Psychology of Learning and Motivation* 36:45–81. [RG]

Shine, J. M., O'Callaghan, C., Halliday, G. M. & Lewis, S. J. G. (2014) Tricks of the mind: Visual hallucinations as disorders of attention. *Progress in Neurobiology* 116:58–65. Available at: http://dx.doi.org/10.1016/j.pneurobio.2014.01.004. [CO]

Shiu, L. & Pashler, H. (1992) Improvement in line orientation discrimination is retinally local but dependent on cognitive set. *Perception and Psychophysics* 52:582–88. [RG]

Siegel, S. (2012) Cognitive penetrability and perceptual justification. *Noûs* 46:201–22. [aCF]

Simons, D. J. (1996) In sight, out of mind: When object representations fail. *Psychological Science* 7:301–305. [aCF]

Simons, D. J. & Chabris, C. F. (1999) Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception* 28:1059–74. [SBM, aCF]

Sinico, M. (2015) Tertiary qualities, from Galileo to Gestalt psychology. *History of the Human Sciences* 28:68–79. [WG]

Sklar, A., Levy, N., Goldstein, A., Mandel, R., Maril, A. & Hassin, R. (2012) Reading and doing arithmetic nonconsciously. *Proceedings of the National Academy of Sciences USA* 109:19614–19. [aCF]

Slotnick, S. D., Schwarzbach, J. & Yantis, S. (2003) Attentional inhibition of visual processing in human striate and extrastriate cortex. *NeuroImage* 19(4):1602–11. [DWV]

Slotnick, S. D., Thompson, W. L., Kosslyn, S. M. (2005) Visual mental imagery induces retinotopically organized activation of early visual areas. *Cerebral Cortex* 15:1570–83. [PDLH]

Sneve, M., Sreenivasan, K., Alaes, D., Endestad, T. & Magnussen, S. (2015) Short-term retention of visual information: Evidence in support of feature-based attention as an underlying mechanism. *Neuropsychologia* 66:1–9. [NB]

Song, H., Vonasch, A. J., Meier, B. P. & Bargh, J. A. (2012) Brighten up: Smiles facilitate perceptual judgment of facial lightness. *Journal of Experimental Social Psychology* 48:450–52. [aCF]

Soudry, Y., Lemogne, C., Malinvaud, D., Consoli, S. M. & Bonfils, P. (2011) Olfactory system and emotion: Common substrates. *European Annals of Otorhinolaryngology, Head and Neck Diseases* 128(1):18–23. [AK]

Spelke, E. S., Breinlinger, K., Jacobson, K. & Phillips, A. (1993) Gestalt relations and object perception: A developmental study. *Perception* 22:1483–501. [GSH]

Sporns, O. (2011) The human connectome: A complex network. *Annals of the New York Academy of Sciences* 1224(1):109–25. [LMH]

Spratling, M. W. (2010) Predictive coding as a model of response properties in cortical area V1. *The Journal of Neuroscience* 30(9):3531–43. [DWV]

Stefanucci, J. K., Gagnon, K. T. & Lessard, D. A. (2011) Follow your heart: Emotion adaptively influences perception. *Social and Personality Psychology Compass* 5:296–308. [aCF]

Stefanucci, J. K., Gagnon, K. T., Tompkins, C. L. & Bullock, K. E. (2012) Plunging into the pool of death: Imagining a dangerous outcome influences distance perception. *Perception* 41:1–11. [aCF]

Stefanucci, J. K. & Geuss, M. N. (2009) Big people, little world: The body influences size perception. *Perception* 38:1782–95. [arCF, LMH]

Stefanucci, J. K. & Proffitt, D. R. (2009) The roles of altitude and fear in the perception of height. *Journal of Experimental Psychology: Human Perception and Performance* 35:424–38. [aCF]

Stefanucci, J. K. & Storbeck, J. J. (2009) Don't look down: Emotional arousal elevates height perception. *Journal of Experimental Psychology: General* 138:131–45. [aCF]

Stein, B. E., Huneycutt, W. S. & Meredith, M. A. (1988) Neurons and behavior: The same rules of multisensory integration apply. *Brain Research* 448(2):355–58. [TE]

Sterling, P. & Laughlin, S. (2015) *Principles of neural design*. MIT Press. [LMH]

Stern, C., Cole, S., Gollwitzer, P., Oettingen, G. & Balcetis, E. (2013) Effects of implementation intentions on anxiety, perceived proximity, and motor performance. *Personality and Social Psychology Bulletin* 39:623–35. [EB]

Sterzer, P., Frith, C. & Petrovic, P. (2008) Believing is seeing: Expectations alter visual awareness. *Current Biology* 18(16):R697–98. Available at: http://doi.org/10.1016/j.cub.2008.06.021. [GL]

Stevenson, R. J. (2009) An initial evaluation of the functions of human olfaction. *Chemical Senses* 35(1):3–20. doi:10.1093/chemse/bjp083. [AK]

Stokes, D. (2013) Cognitive penetrability of perception. *Philosophy Compass* 8:646–63. [aCF]

Stokes, D. (2014) Cognitive penetration and the perception of art. *Dialectica* 68:1–34. [aCF]

Stokes, M., Thompson, R., Cusack, R. & Duncan, J. (2009) Top-down activation of shape-specific population codes in visual cortex during mental imagery. *Journal of Neuroscience* 29:1565–72. [PDLH]

Storbeck, J. & Clore, G. L. (2008) Affective arousal as information: How affective arousal influences judgments, learning, and memory. *Social and Personality Psychology Compass* 2:1824–43. [aCF]

Storbeck, J. & Stefanucci, J. K. (2014) Conditions under which arousal does and does not elevate height estimates. *PLoS ONE* 9:e92024. [aCF]

Sugovic, M. & Witt, J. K. (2013) An older view on distance perception: Older adults perceive walkable extents as farther. *Experimental Brain Research* 226:383–91. [aCF]

Suzuki, S. & Cavanagh, P. (1997) Focused attention distorts visual space: An attentional repulsion effect. *Journal of Experimental Psychology: Human Perception and Performance* 23(2):443–63. [BdH]

Taylor, J. E. T., Witt, J. K. & Sugovic, M. (2011) When walls are no longer barriers: Perception of wall height in parkour. *Perception* 40:757–60. [aCF]

Taylor-Covill, G. A. H. & Eves, F. F. (2016) Carrying a biological "backpack": Quasi-experimental effects of weight status and body fat change on perceived steepness. *Journal of Experimental Psychology: Human Perception and Performance* 42(3):331–38. [GLC, rCF]

Teachman, B. A., Stefanucci, J. K., Clerkin, E. M., Cody, M. W. & Proffitt, D. R. (2008) A new mode of fear expression: Perceptual bias in height fear. *Emotion* 8:296–301. [aCF]

Teufel, C., Subramaniam, N., Dobler, V., Perez, J., Finnemann, J., Mehta, P. R., Goodyer, I. M. & Fletcher, P. C. (2015) Shift toward prior knowledge confers a perceptual advantage in early psychosis and psychosis-prone healthy individuals. *Proceedings of the National Academy of Sciences USA* 112(43):13401–406. [CO]

Theeuwes, J. (2013) Feature-based attention: It is all bottom-up priming. *Philosophical Transactions of the Royal Society B* 368(1628):1–11. [NB]

Thomas, G. (1941) Experimental study of the influence of vision on sound localization. *Journal of Experimental Psychology* 28(2):163–77. [BdH]

Toccafondi, F. (2009) Facts, values, emotions, and perception. In: *Values and ontology: Problems and perspectives*, ed. B. Centi & W. Huemer, pp. 137–54. Ontos Verlag. [WG]

Toppino, T. C. (2003) Reversible-figure perception: Mechanisms of intentional control. *Perception and Psychophysics* 65:1285–95. [arCF]

Tranel, D., Damasio, H. & Damasio, A. R. (1995) Double dissociation between overt and covert face recognition. *Journal of Cognitive Neuroscience* 7(4):425–32. [RMR]

Trapp, S. & Bar, M. (2015) Prediction, context and competition in visual recognition. *Annals of the New York Academy of Sciences* 1339:190–98. doi:10.1111/nyas.12680. [CO]

Tse, P. U. (2005) Voluntary attention modulates the brightness of overlapping transparent surfaces. *Vision Research* 45:1095–98. [aCF]

Tse, P. U. & Cavanagh, P. (2000) Chinese and Americans see opposite apparent motions in a Chinese character. *Cognition* 74:B27–32. [aCF]

Tseng, P. & Bridgeman, B. (2011) Improved change detection with nearby hands. *Experimental Brain Research* 209:257–69. [PT]

Tseng, P., Bridgeman, B. & Juan, C. H. (2012) Take the matter into your own hands: A brief review of the effect of nearby-hands on visual processing. *Vision Research* 72:74–77. [PT]

Tubridy, S. & Davachi, L. (2011) Medial temporal lobe contributions to episodic sequence encoding. *Cerebral Cortex* 21(2):272–80. doi:10.1093/cercor/bhq092. [LLE]

Turk-Browne, N. B., Junge, J. A. & Scholl, B. J. (2005) The automaticity of visual statistical learning. *Journal of Experimental Psychology: General* 134:552–64. [rCF]

Turvey, M. T., Shaw, R. E., Reed, E. S. & Mace, W. M. (1981) Ecological laws of perceiving and acting: In reply to Fodor and Pylyshyn (1981). *Cognition* 9 (3):237–304. [PT]

Uncapher, M. R. & Rugg, M. D. (2009) Selecting for memory? The influence of selective attention on the mnemonic binding of contextual information. *The Journal of Neuroscience* 29(25):8270–79. [DWV]

Ungerleider, L. & Mishkin, M. (1982) Two cortical visual systems. In: *Analysis of visual behavior*, ed. D. Ingle, M. Goodale & R. Mansfield, pp. 549–86. The MIT Press. [CO]

Van Bavel, J. J., FeldmanHall, O. & Mende-Siedlecki, P. (2015) The neuroscience of moral cognition: From dual processes to dynamic systems. *Current Opinion in Psychology* 6:167–72. [APG]

van den Heuvel, M. P. & Sporns, O. (2013) Network hubs in the human brain. *Trends in Cognitive Sciences* 17(12):683–96. [LMH]

van der Kamp, J., Rivas, F., van Doorn, H. & Savelsbergh, G. J. P. (2008) Ventral and dorsal contributions in visual anticipation in fast ball sports. *International Journal of Sport Psychology* 39:100–30. [RC-B]

van Koningsbruggen, G. M., Stroebe, W. & Aarts, H. (2011) Through the eyes of dieters: Biased size perception of food following tempting food primes. *Journal of Experimental Social Psychology* 47:293–99. [aCF]

van Loon, A. M., Fahrenfort, J. J., van der Velde, B., Lirk, P. B., Vulink, N. C. C., Hollmann, M. W., Scholte, S. & Lamme, V. A. F. (2015) NMDA receptor antagonist ketamine distorts object recognition by reducing feedback to early visual cortex. *Cerebral Cortex* 26(5):1986–96. . Available at: http://doi.org/10.1093/cercor/bhv018. [DMB]

van Ulzen, N. R., Semin, G. R., Oudejans, R. R. D. & Beek, P. J. (2008) Affective stimulus properties influence size perception and the Ebbinghaus illusion. *Psychological Research* 72:304–10. [aCF]

Vandenbroucke, A. R., Fahrenfort, J. J., Meuwese, J. D., Scholte, H. S. & Lamme, V. A. (2014) Prior knowledge about objects determines neural color representation in human visual cortex. *Cerebral Cortex* 26(4):1401–408. [CW]

Verducci, T. (2000) The power of Pedro. *Sports Illustrated* [vol(issue)]:, 54–62. Available at: http://www.si.com/mlb/2015/07/23/pedro-martinez-hall-of-fame-boston-red-sox-tom-verducci, May 18, 2016, 54–62. [JKW]

Vetter, P. & Newen, A. (2014) Varieties of cognitive penetration in visual perception. *Consciousness and Cognition* 27:62–75. [NB, PDLH, aCF]

Vetter, P., Smith, F. W. & Muckli, L. (2014) Decoding sound and imagery content in early visual cortex. *Current Biology* 24(11):1256–62. doi:10.1016/j.cub.2014.04.020. [BdH]

Vickery, T. J., Chun, M. M. & Lee, D. (2011) Ubiquity and specificity of reinforcement signals throughout the human brain. *Neuron* 72:166–77. [aCF]

Vollenweider, F. X. (2001) Brain mechanisms of hallucinogens and entactogens. *Dialogues in Clinical Neuroscience* 3:265–79. [PDLH]

Vossel, S., Mathys, C., Daunizeau, J., Bauer, M., Driver, J., Friston, K. J. & Stephan, K. E. (2014) Spatial attention, precision, and Bayesian inference: A study of saccadic response speed. *Cerebral Cortex* 24:1436–50. [ACl]

Vrins, S., de Wit, T. C. J. & van Lier, R. (2009) Bricks, butter, and slices of cucumber: Investigating semantic influences in amodal completion. *Perception* 38(1):17–29. [GL]

Vukicevic, M. & Fitzmaurice, K. (2008) Butterflies and black lacy patterns: The prevalence and characteristics of Charles Bonnet hallucinations in an Australian population. *Clinical and Experimental Ophthalmology* 36:659–65. [PDLH]

von Helmholtz, H. (2005) *Treatise on physiological optics, vol. 3*. Courier Dover. [TE]

Wakslak, C. J. & Kim, B. K. (2015) Controllable objects seem closer. *Journal of Experimental Psychology: General* 144:522–27. [aCF]

Walther, D. B., Caddigan, E., Fei-Fei, L. & Beck, D. M. (2009) Natural scene categories revealed in distributed patterns of activity in the human brain. *Journal of Neuroscience* 29(34):10573–81. Available at: http://doi.org/10.1523/JNEUROSCI.0559-09.2009. [DMB]

Wang, L., Kennedy, B. L. & Most, S. B. (2012) When emotion blinds: A spatiotemporal competition account of emotion-induced blindness. *Frontiers in Psychology: Special Topic on Emotion and Cognition* 3:438. [SBM]

Ward, E. J. & Scholl, B. J. (2015) Inattentional blindness reflects limitations on perception, not memory: Evidence from repeated failures of awareness. *Psychonomic Bulletin and Review* 22:722–27. [SBM, arCF]

Ward, J. & Meijer, P. (2010) Visual experiences in the blind induced by an auditory sensory substitution device. *Consciousness and Cognition* 19(1):492–500. doi:10.1016/j.concog.2009.10.006. [TE]

Warren, W. H. (1984) Perceiving affordances: Visual guidance of stair climbing. *Journal of Experimental Psychology: Human Perception and Performance* 10:683–703. [rCF]

Waters, F., Collerton, D., ffytche, D. H., Jardri, R., Pins, D., Dudley, R., Blom, J. D., Mosimann, U. P., Eperjesi, F. & Ford, S. (2014) Visual hallucinations in the psychosis spectrum and comparative information from neurodegenerative disorders and eye disease. *Schizophrenia Bulletin* 40(Suppl 4):S233–45 [CO]

Watkins, S., Shams, L., Tanaka, S., Haynes, J.-D. & Rees, G. (2006) Sound alters activity in human V1 in association with illusory visual perception. *NeuroImage* 31(3):1247–56. doi:10.1016/j.neuroimage.2006.01.016. [BdH]

Weber, A. (2001) Help or hindrance: How violation of different assimilation rules affects spoken-language processing. *Language and Speech* 44:95–118. [ACu]

Webster, M. A. & Kay, P. (2012) Color categories and color appearance. *Cognition* 122:375–92. [aCF]

Weinberger, N. M. (2004) Specific long-term memory traces in primary auditory cortex. *Nature Reviews Neuroscience* 5:279–90. [VM]

Wesp, R., Cichello, P., Gracia, E. B. & Davis, K. (2004) Observing and engaging in purposeful actions with objects influences estimates of their size. *Perception and Psychophysics* 66:1261–67. [arCF]

Wesp, R. & Gasper, J. (2012) Is size misperception of targets simply justification for poor performance? *Perception* 41:994–96. [arCF]

White, E. J., Shockley, K. & Riley, M. A. (2013) Multimodally specified energy expenditure and perceived distance. *Psychonomic Bulletin and Review* 20:1371–77. [GLC]

Witt, J. K. (2011a) Action's effect on perception. *Current Directions in Psychological Science* 20:201–206. [aCF]

Witt, J. K. (2011b) Tool use influences perceived shape and perceived parallelism, which serve as indirect measures of perceived distance. *Journal of Experimental Psychology: Human Perception and Performance* 37:1148–56. [aCF]

Witt, J. K. & Dorsch, T. E. (2009) Kicking to bigger uprights: Field goal kicking performance influences perceived size. *Perception* 38:1328–40. [aCF]

Witt, J. K., Linkenauger, S. A., Bakdash, J. Z. & Proffitt, D. R. (2008) Putting to a bigger hole: Golf performance relates to perceived size. *Psychonomic Bulletin and Review* 15:581–85. [aCF]

Witt, J. K. & Proffitt, D. R. (2005) See the ball, hit the ball. *Psychological Science* 16:937–38. [aCF]

Witt, J. K., Proffitt, D. R. & Epstein, W. (2004) Perceiving distance: A role of effort and intent. *Perception* 33:577–90. [arCF]

Witt, J. K., Proffitt, D. R. & Epstein, W. (2005) Tool use affects perceived distance, but only when you intend to use it. *Journal of Experimental Psychology: Human Perception and Performance* 31:880–88. [aCF]

Witt, J. K., Proffitt, D. R. & Epstein, W. (2010) When and how are spatial perceptions scaled? *Journal of Experimental Psychology: Human Perception and Performance* 36:1153–60. [aCF]

Witt, J. K. & Sugovic, M. (2010) Performance and ease influence perceived speed. *Perception* 39(10):1341–53. doi:10.1068/P6699. [JKW, arCF]

Witt, J. K. & Sugovic, M. (2012) Does ease to block a ball affect perceived ball speed? Examination of alternative hypotheses. *Journal of Experimental Psychology: Human Perception and Performance* 38(5):1202–14. doi:10.1037/a0026512. [JKW]

Witt, J. K. & Sugovic, M. (2013a) Catching ease influences perceived speed: Evidence for action-specific effects from action-based measures. *Psychonomic Bulletin and Review* 20:1364–70. [JKW]

Witt, J. K. & Sugovic, M. (2013b) Response bias cannot explain action-specific effects: Evidence from compliant and non-compliant participants. *Perception* 42:138–52. [JKW, rCF]

Witt, J. K., Schuck, D. M. & Taylor, J. E. T. (2011) Action-specific effects underwater. *Perception* 40:530–37. [aCF]

Witt, J. K., Sugovic, M. & Dodd, M. D. (2016) Action-specific perception of speed is independent of attention. *Attention, Perception, & Psychophysics* 78(3):880–90. doi:10.3758/s13414-015-1047-6 [JKW]

Witt, J. K., Sugovic, M. & Taylor, J. E. T. (2012) Action-specific effects in a social context: Others' abilities influence perceived speed. *Journal of Experimental Psychology: Human Perception and Performance* 38(3):715–25. doi:10.1037/a0026261. [JKW]

Witzel, C. & Hansen, T. (2015) Memory effects on colour perception. In: *Handbook of color psychology*, ed. A. J. Elliot, M. D. Fairchild & A. Franklin, pp. 641–65. Cambridge University Press. Available at: http://ebooks.cambridge.org/chapter.jsf?bid=CBO9781107337930&cid=CBO9781107337930A [CW]

Witzel, C., Valkova, H., Hansen, T. & Gegenfurtner, K. R. (2011) Object knowledge modulates colour appearance. *i-Perception* 2(1):13–49. [CW, aCF]

Wolfe, J. M. (1999) Inattentional amnesia. In: *Fleeting memories: Cognition of brief visual stimuli*, ed. V. Coltheart, pp. 71–94. MIT Press. [SBM]

Womelsdorf, T., Anton-Erxleben, K. & Treue, S. (2008) Receptive field shift and shrinkage in macaque middle temporal area through attentional gain modulation. *The Journal of Neuroscience* 28(36):8934–44. doi:10.1523/JNEUROSCI.4030-07.2008. [BdH]

Woods, A. J., Philbeck, J. W. & Danoff, J. V. (2009) The various perceptions of distance: An alternative view of how effort affects distance judgments. *Journal of Experimental Psychology: Human Perception and Performance* 35:1104–17. [aCF]

Wunderlich, K., Schneider, K. A. & Kastner, S. (2005) Neural correlates of binocular rivalry in the human lateral geniculate nucleus. *Nature Neuroscience* 8 (11):1595–602. Available at: http://doi.org/10.1038/nn1554. [DMB]

Wündt, W. (1897) *Grundriss der Psychologie, von Wilhelm Wundt*. W. Engelmann. [VM]

Wyatte, D., Curran, T. & O'Reilly, R. (2012) The limits of feedforward vision: Recurrent processing promotes robust object recognition when objects are degraded. *Journal of Cognitive Neuroscience* 24:2248–61. [RG]

Xie, W. & Zhang, W. (in preparation) Positive thought as an eye opener leads to brighter perception. [WX]

Yantis, S. (2013) *Sensation and perception*. Palgrave Macmillan. [aCF, TE]

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C. & Wager, T. D. (2011) Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods* 8(8):665–70. [LMH]

Yee, E., Ahmed, S. Z. & Thompson-Schill, S. L. (2012) Colorless green ideas (can) prime furiously. *Psychological Science* 23:364–69. [aCF]

Yeo, B. T., Krienen, F. M., Eickhoff, S. B., Yaakub, S. N., Fox, P. T., Buckner, R. L., Asplund, C. L. & Chee, M. W. (2015) Functional specialization and flexibility in human association cortex. *Cerebral Cortex* 25(10):3654–72. [LMH]

Zadra, J. R. & Clore, G. L. (2011) Emotion and perception: The role of affective information. *Wiley Interdisciplinary Reviews: Cognitive Science* 2:676–85. [aCF]

Zadra, J. R., Weltman, A. L. & Proffitt, D. R. (2016) Walkable distances are bioenergetically scaled. *Journal of Experimental Psychology: Human Perception and Performance* 42(1): 39–51. Available at: http://dx.doi.org/10.1037/xhp0000107. [GLC]

Zanto, T. P. & Gazzaley, A. (2009) Neural suppression of irrelevant information underlies optimal working memory performance. *The Journal of Neuroscience* 29(10):3059–66. [DWV]

Zeimbekis, J. (2013) Color and cognitive penetrability. *Philosophical Studies* 165:167–75. [rCF]

Zeimbekis, J. & Raftopoulos, A., eds. (2015) *Cognitive effects on perception: New philosophical perspectives*. Oxford University Press. [AR, aCF]

Zhang, P., Jamison, K., Engel, S., He, B. & He, S. (2011) Binocular rivalry requires visual attention. *Neuron* 71(2):362–69. [DWV]

Zhang, S., Xu, M., Kamigaki, T., Hoang Do, J. P., Chang, W. C. & Jenvay, S. et al. (2014) Long-range and local circuits for top-down modulation of visual cortex processing. *Science* 345:660–65. [aCF]

Zhang, W. & Luck, S. J. (2009) Feature-based attention modulates feedforward visual processing. *Nature Neuroscience* 12(1):24–25. Available at: http://doi.org/10.1038/nn.2223. [WX]