

Below is an unedited, uncorrected BBS Target Article recently accepted for publication. This preprint has been prepared specifically for potential commentators who wish to nominate themselves for formal commentary invitation via Editorial Manager: <http://bbs.edmgr.com/>. The Commentary Proposal Instructions can be accessed here: <http://journals.cambridge.org/action/displaySpecialPage?pageId=5544>

Please DO NOT write a commentary unless you receive a formal email invitation from the Editors. If you are invited to submit a commentary, a copyedited, corrected version of this paper will be made available.

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

Chaz Firestone and Brian Scholl
Yale University

Correspondence:
Brian Scholl
Department of Psychology, Yale University, Box 208205, New Haven, CT 06520-8205
203-432-4629
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 — provide 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 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.

Keywords: Top-down effects; Modularity; Encapsulation; Cognitive penetration; Perception; Cognition

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 this 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 one hand, *seeing* a red apple and, on the other hand, *thinking about*, *remembering* or *desiring* a red apple. This 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 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, 2015; 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.

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 160 papers published since 1995 reporting such effects; for a list, see <http://www.yale.edu/perception/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 morally relevant words 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 such factors do influence how we see, then such models and textbooks are scandalously incomplete.

While such factors 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 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). In an especially sweeping claim, this sort of evidence led one philosopher to declare 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, 2001b) 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 importantly, 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 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. These pitfalls therefore must be taken seriously by research on top-down effects on visual perception 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 thus no surprise that this 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” which directly addresses this aspect of how the mind is organized. This 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, etc. We single out this meaning of “top-down” not only because it may be the most prominent usage of this 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 vs. (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

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

organized — and for an obvious reason: closing your eyes (or turning off the lights) changes only the *input* to perception, without changing perceptual processing itself.

Despite the triviality of this example, it is worth keeping this distinction in mind, since it is not always obvious when an effect operates by changing the input. To take one fascinating example, it has recently been shown that facial expressions associated with fear (e.g. widened eyes) and disgust (e.g. narrowed eyes) 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 in regards 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, via 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 & Wu, 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 Section 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 (for example, see Raftopoulos, 2001a). 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 vs. 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, these 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 these 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 Figure 1). In such cases, the underlying processes continue to operate reflexively (based only 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 (Figure 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 disks 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 too — 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 cross-modal 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. under 100ms; 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 reviewed below 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; Figure 2a). However, such changes don't involve any penetration since 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). Thus, what may superficially appear to be an influence of knowledge on perception is simply increased sensitivity to certain contours. Indeed, the authors of these reports themselves 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" (Peterson & Gibson, 1994, p.561). Thus, such effects aren't "top-down" in any sense that implies cognitive penetrability, since 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.)

3. Contemporary Top-Down Effects

What remains after setting aside alternate 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. P.M. Churchland, 1988; P.S. Churchland et al., 1994; Firestone, 2013a; Fodor, 1983, 1984, 1988; Goldstone & Barsalou, 1998; Lupyan, 2012; Machery, in press; Proffitt & Linkenauger, 2013; Pylyshyn, 1999; Raftopolous, 2001a; Raftopolous & Zeimbekis, in press; Vetter & Newen, 2014). 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 judgments were made 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 non-uniform, 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 (Balci et al., 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; Figure 2b). Morally relevant words reportedly "pop out" in visual awareness when briefly presented (Gantman & Van Bavel, 2014; Figure 2c), and follow-up studies suggested that this 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; Figure 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; Figure 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, in press); subjects who drink a sugary beverage (rather than a low-calorie alternative) estimate hills as shallower (Schnall et al., 2010); and swimmers who wear speedy 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 uprights (Witt & Dorsch, 2009), lower tennis nets (Witt & Sugovic, 2010), larger dart-boards (Cañal-Bruland et al., 2010; Wesp et al., 2004; Figure 2f), larger golf holes (Witt et al., 2008), and (for *parkour* experts) shorter walls (Taylor et al., 2010). 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 itself 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, there are recent reports that thinking negative thoughts makes the world look darker (Banerjee et al., 2012; Meier et al., 2007; Figure 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 alligator/squirrel) take on their scarier interpretations (Prinz & Seidel, 2012; Figure 2h); social exclusion makes other people look closer (Pitts et al., 2014); and smiling faces appear brighter (Song et al., 2012; Figure 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; Figure 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. Figure 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 (Figure 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 has been 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; Figure 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; Figure 2n).

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

4. The 6 '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 are empirically grounded — not just in the weak sense that they discuss relevant empirical evidence, but in the stronger sense that they have been demonstrated to explain 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. Finally, (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 non-revolutionary — 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 you should also *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, 2014) 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 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 El Greco’s 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; Figure 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 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, 2014, 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 (Figure 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 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 in fact this experiment did not 'fail': subjects again reported narrower apertures even when responding with an aperture (Firestone & Scholl, 2014, Experiment 2). Thus, 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, non-perceptual, 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, or grayscale patches) and the item reproduced (e.g. visual scenes to be painted, or 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; Figure 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 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; Figure 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 should have canceled out.

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 pro-actively employ both confirmatory and disconfirmatory research strategies; to do otherwise is to ignore half 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 via positive replications. Second, the El Greco strategy can fuel such implications even before researchers determine the actual (non-perceptual) 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 anytime

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 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.

a scale can be influenced just as the critical stimuli are supposedly influenced (e.g. nearly all perceptual matching tasks).

4.2 Pitfall #2: Perception vs. 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 literally 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 (but 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 also conclude or infer that an object is red. Thus, any time an experiment shifts perceptual reports, it is possible that this 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 — but 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 how the distance truly *looked*.

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; Figure 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, this 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 readymade 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., Balci et al., 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 & Balci et al., 2011) or one’s work office (Wakslak & Kim, in press). 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 this effect disappeared when subjects made their throws while perched atop a 12-foot-high rock-climbing wall (Cañal-Bruland et al., 2010). Though this was interpreted as an effect of anxiety on action-specific perception, this finding could easily be recast as an effect on judgment: 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 — since 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, yet 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 — e.g. 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 this distinction, it seems incumbent on any proposal of a top-down effect to explicitly and prominently address this distinction, if only rhetorically — since 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 — and this raises the possibility that social biases may intrude on perceptual reports in a more specific way than the previous pitfall. Whereas *judgments* of various visual qualities are often sincerely held even when they are subject to top-down influence (such that, for example, 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 reports (of anything, including percepts) to be 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 and Rubin, 1978.)

Contamination by demand characteristics seems especially likely in experiments involving a single conspicuous manipulation and a single perceptual report. But even moreso

than the previous pitfall, it seems especially easy to combat such influences — either 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; Figure 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, 2014) 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 — e.g. 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; Figure 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 these biases for the worthless \$0 gift card, which wouldn’t have been worth any such strategizing. A follow-up study (Durgin et al., 2011) 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 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 only 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 ‘correct’ 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; Figure 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 the manner of backpacks and hills (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 non-scary ambiguous figures (Prinz & Seidel, 2012; Figure 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, since assessing demand effects is often so easy and cost-free. In particular, although demand effects can be mitigated by non-transparent manipulations or indirect measures, they can also often be assessed by simply *asking the subjects about the experiment* — e.g. during a careful post-experimental debriefing. For example, before the experiment's purpose is revealed, subjects can be carefully asked what they thought the experiment was about, what strategies they used, etc. These sorts of questions can readily reveal (or help to rule out) active demand factors.

Such debriefing was especially helpful, for example, in the case of backpacks and reported slant, where 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 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. their 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 vs. non-arousing 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 (as live tarantulas might differ in size, color, and motion from plush toys) — and that these low-level differences are actually responsible for perceptual differences across conditions.

We have suggested (and will continue to suggest) that many pitfalls discussed 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.

This 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 in often-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, this 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 may often directly rule out the kinds of worries expressed in the previous two pitfalls (e.g. task demands and post-perceptual 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 over 20% yellow (Hansen et al., 2006; Olkonen et al., 2008); however, if you look now at a grayscale image of a banana (Figure 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 good vs. bad words and smiling vs. frowning faces. This is because when the effects weren't caricatured in this way, readers could not understand the claims, since 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 Figure 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, while the White face has darker eyes. Of course, there must exist *some* low-level differences between the images, since 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 while blurring 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 effectively shows how the lightness difference can derive from low-level visual features — which,

critically, are present in the original images — without any contribution by perceived race. (And note that such results are unlikely to reflect *unconscious* race categorization — since it would be a distinctively odd implicit race judgment that could influence explicit lightness judgments *but not explicit race judgments*.) And while 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 were initially seen as reflecting high-level factors have been shown to reflect lower-level visual differences across conditions. For example, it was reported that visual search with line drawings of animals and artifacts is facilitated by categorical differences — e.g. with faster and more efficient searches for animals among artifacts, and vice versa (Levin et al., 2001). This initially appeared to be a high-level effect on a fairly low-level perceptual process, given that efficient visual search is often considered ‘preattentive’. However, on closer investigation (and to their immense credit), these researchers themselves 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 followup 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 — as when a central dog image looks smaller when surrounded by images of larger dogs vs.

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, since 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, since they got the effects even when reporting the lightness of a face *using a copy of that exact face*).

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; Figure 21), 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, spiders have been contrasted not with plush toys, but with ‘scrambled’ spiders, or even with 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, since 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 which carry no such implications (see Section 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 Section 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., 2005; 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).

These 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 this happens — say, if we attend to a specific flower and it looks clearer or more detailed — should this 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., 2005; 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 *why* you attended that way. Having the lights turned off will have the same effect on visual perception regardless of why they were

turned 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 is seen regardless of the reasons that led you to deploy attention in that way, and even whether you attended voluntarily or via 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.

Due to 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 one hand, perhaps based on its prominence in previous discussions (especially Pylyshyn, 1999), the kinds of thoughts from 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 via 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 the first place 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 — especially given that such effects are often straightforwardly testable. In particular, for a possible influence of intention on perception (for example), it is almost always possible to separate intention from attention — e.g. by holding one constant while varying the other. For example, to factor attention out, one can impose an attentional load, often via 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 (Figure 2r) reveals that one can voluntarily switch which interpretation is seen (in particular, which face of the cube seems to be in front), and this 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 have suggested 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, in press; 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 relatively 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 biases the interpretation of ambiguous figures (such as an alligator / squirrel figure; Figure 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") will 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 and 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; Figure 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 instance, 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 visibly darkens transparent surfaces (Tse, 2005), which could explain the increase in perceived transparency among thirsty observers (Changizi & Hall, 2001; Figure 2b) if non-thirsty 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 since 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 Section 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 could 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 semi-transparent surface does not make it look more opaque, then this could eliminate 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 little 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*, yet 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 recognitional tasks (Lupyan & Spivey, 2008; Lupyan et al., 2010; Figure 2m), and that, when briefly presented, morally relevant words are easier to identify than morally irrelevant words (the “moral popout effect” as in Figure 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: in order 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, since 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 popout 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 presented words and non-words (40-60ms) one at a time, and decided whether each presented stimulus was a word or a non-word (see Figure 4a). Some words were morally relevant (e.g. “illegal”), and some were morally irrelevant (e.g. “limited”), and 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* (for example, 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.

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), where 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 “popout effect” in similar circumstances — including completely arbitrary categories without any special importance in the mind. To test this possibility, we replicated the “moral popout effect” (Firestone & Scholl, 2015b), but instead of morally related words (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 Figure 4b). (A second experiment replicated the phenomenon again, with transportation-related words such as “car” and “passenger”.) This suggests that relatedness is the key factor in such effects, and thus that memory, not perception, improves detection of morally related words. 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 squiggle symbol (\sqcap) appeared as an oddball item in a search array populated by mirror images of that symbol (\sqcup), subjects who were told to think of the symbols as resembling “2” and “5” were faster at finding a \sqcap among \sqcup s than were naïve 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 \square among \square 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 was 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 may 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 non-hungry subjects (Radel and Clément-Guillotin, 2012); but if hungry subjects were simply *thinking about food* more than non-hungry 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 itself would already 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 those mentioned in Section 4.6) do not even acknowledge that such an interpretation is important or interesting. (Lupyan and 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 so; 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 non-semantic 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 so 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 9 times that many empirical reports of top-down effects (and of course that only includes 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 coin-size paper 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”: to 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 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 most of these pitfalls were never considered 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 — e.g. in situations characterized by an “El Greco fallacy”.
2. Perception vs. Judgment: Disentangle post-perceptual judgment from actual online perception — e.g. 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 — e.g. 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 imposing an attentional load to attenuate such influences, examine whether higher-level states directly influence lower-level visual processes, or instead whether 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, e.g. 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 possibilities. However, such a checklist can be usefully employed simply by taking care to explicitly discuss (or at least mention!) each potential alternate 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 simply a demonstration that people pay attention to objects they like.

The possibility of top-down effects on perception is tremendously exciting, and 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.

References

- Abrams, R. A., & Weidler, B. J. (in press). How far away is that? It depends on you: Perception accounts for the abilities of others. *Journal of Experimental Psychology: Human Perception and Performance*.
- Adelson, E. H. (2000). Lightness perception and lightness illusions. In M. Gazzaniga (Ed.), *The New Cognitive Neurosciences* (2nd ed., pp. 339–351). Cambridge, MA: MIT Press.
- 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.
- Anstis, S. (2002). Was El Greco astigmatic? *Leonardo, 35*, 208.
- Anton-Erxleben, K., Henrich, C., & Treue, S. (2007). Attention changes perceived size of moving visual patterns. *Journal of Vision, 7*, 5:1–9.
- Balcetis, E. (2015). Approach and avoidance as organizing structures for motivated distance perception. *Emotion Review, 9*, 1096–1096.
- Balcetis, E., & Dunning, D. (2006). See what you want to see: Motivational influences on visual perception. *Journal of Personality and Social Psychology, 91*, 612–625.
- Balcetis, E., & Dunning, D. (2010). Wishful seeing: More desired objects are seen as closer. *Psychological Science, 21*, 147–152.
- Balcetis, E., Dunning, D., & Granot, Y. (2012). Subjective value determines initial dominance in binocular rivalry. *Journal of Experimental Social Psychology, 48*, 122–129.
- Banerjee, P., Chatterjee, P., & Sinha, J. (2012). Is it light or dark? Recalling moral behavior changes perception of brightness. *Psychological Science, 23*, 407–409.
- Bannert, M. M., & Bartels, A. (2013). Decoding the yellow of a gray banana. *Current Biology, 23*, 2268–2272.
- Bhalla, M., & Proffitt, D. R. (1999). Visual-motor recalibration in geographical slant perception. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1076–1096.

- Blake, R., & Sekuler, R. (2005). *Perception* (5th ed.). New York: McGraw-Hill.
- Bouvier, S., & Treisman, A. (2010). Visual feature binding requires reentry. *Psychological Science*, *21*, 200–204.
- Bruner, J. S. (1957). On perceptual readiness. *Psychological Review*, *64*, 123–152.
- Bruner, J. S., Postman, L., & Rodrigues, J. (1951). Expectation and the perception of color. *American Journal of Psychology*, *64*, 216–227.
- Bruner, J. S., & Goodman, C. C. (1947). Value and need as organizing factors in perception. *Journal of Abnormal and Social Psychology*, *42*, 33–44.
- Cañal-Bruland, R., Pijpers, J. R. R., & Oudejans, R. R. D. (2010). The influence of anxiety on action-specific perception. *Anxiety, Stress & Coping*, *23*, 353–361.
- 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*, 285–289.
- Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, *7*, 308–313.
- Carter, L. F., & Schooler, K. (1949). Value, need, and other factors in perception. *Psychological Review*, *56*, 200–207.
- 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*, *106*, 20168–20173.
- Cave, K. R., & Bichot, N. P. (1999). Visuospatial attention: Beyond a spotlight model. *Psychonomic Bulletin & Review*, *6*, 204–223.
- Changizi, M. A., & Hall, W. G. (2001). Thirst modulates a perception. *Perception*, *30*, 1489–1497.
- 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.
- Churchland, P. M. (1988). Perceptual plasticity and theoretical neutrality: A reply to Jerry Fodor.

Philosophy of Science, 55, 167–187.

Churchland, P. S., Ramachandran, V. S., & Sejnowski, T. J. (1994). A critique of pure vision. In C. Koch & J. Davis (Eds.), *Large-Scale Neuronal Theories of the Brain* (pp. 23–60). Cambridge, MA: MIT Press.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36, 181–204.

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, 381–386.

Cohen, M. A., Alvarez, G. A., & Nakayama, K. (2011). Natural-scene perception requires attention. *Psychological Science*, 22, 1165–1172.

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

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

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82, 407–428.

Collins, J. A., & Olson, I. R. (2014). Knowledge is power: How conceptual knowledge transforms visual cognition. *Psychonomic Bulletin & Review*, 21, 843–860.

Cooper, A. D., Sterling, C. P., Bacon, M. P., & Bridgeman, B. (2012). Does action affect perception or memory? *Vision Research*, 62, 235–240.

Coren, S., & Enns, J. T. (1993). Size contrast as a function of conceptual similarity between test and inducers. *Perception & Psychophysics*, 54, 579–588.

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, 219–233.

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, 362–368.

Dils, A. T., & Boroditsky, L. (2010a). Processing unrelated language can change what you see.

Psychonomic Bulletin & Review, 17, 882–888.

Dils, A. T., & Boroditsky, L. (2010b). Visual motion aftereffect from understanding motion

language. *Proceedings of the National Academy of Sciences*, 107, 16396–16400.

Doerrfeld, A., Sebanz, N., & Shiffrar, M. (2012). Expecting to lift a box together makes the load

look lighter. *Psychological Research*, 76, 467–475.

Dunning, D., & Balcetis, E. (2013). Wishful seeing: How preferences shape visual perception.

Current Directions in Psychological Science, 22, 33–37.

Durgin, F. H., Baird, J. A., Greenburg, M., Russell, R., Shaughnessy, K., & Waymouth, S. (2009).

Who is being deceived? The experimental demands of wearing a backpack. *Psychonomic*

Bulletin & Review, 16, 964–969.

Durgin, F. H., DeWald, D., Lechich, S., Li, Z., & Ontiveros, Z. (2011). Action and motivation:

Measuring perception or strategies? *Psychonomic Bulletin & Review*, 18, 1077–1082.

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, 1582–1595.

Erdelyi, M. (1974). A new look at the new look: Perceptual defense and vigilance. *Psychological*

Review, 81, 1–25.

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–

473.

Firestone, C. (2013b). On the origin and status of the "El Greco fallacy". *Perception*, 42, 672–674.

Firestone, C., & Scholl, B. J. (2014). "Top-down" effects where none should be found: The El

Greco fallacy in perception research. *Psychological Science*, 25, 38–46.

Firestone, C., & Scholl, B. J. (2015a). Can you *experience* top-down effects on perception? The

- case of race categories and perceived lightness. *Psychonomic Bulletin & Review*, 22, 694–700.
- Firestone, C., & Scholl, B. J. (2015b). Enhanced visual awareness for morality and pajamas? Perception vs. memory in top-down effects. *Cognition*, 136, 409–416.
- Fodor, J. A. (1983). *The modularity of mind: An essay in faculty psychology*. Cambridge, MA: MIT Press.
- Fodor, J. A. (1984). Observation reconsidered. *Philosophy of Science*, 51, 23–43.
- Fodor, J. A. (1988). A reply to Churchland's "Perceptual plasticity and theoretical neutrality". *Philosophy of Science*, 55, 188–198.
- Gantman, A. P., & Van Bavel, J. J. (2014). The moral pop-out effect: Enhanced perceptual awareness of morally relevant stimuli. *Cognition*, 132, 22–29.
- Gao, T., McCarthy, G., & Scholl, B. J. (2010). The wolfpack effect: Perception of animacy irresistibly influences interactive behavior. *Psychological Science*, 21, 1845–1853.
- 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 & Psychophysics*, 72, 1890–1902.
- Gibson, J. J. (1979). *The ecological approach to perception and action*. Boston, MA: Houghton Mifflin.
- Gilbert, C. D., & Li, W. (2013). Top-down influences on visual processing. *Nature Reviews Neuroscience*, 14, 350–363.
- Gilchrist, A., & Jacobsen, A. (1984). Perception of lightness and illumination in a world of one reflectance. *Perception*, 13, 5–19.
- Gobell, J., & Carrasco, M. (2005). Attention alters the appearance of spatial frequency and gap size. *Psychological Science*, 16, 644–651.
- Goldstone, R. L. (1995). Effects of categorization on color perception. *Psychological Science*, 6, 298–304.
- Goldstone, R. L., & Barsalou, L. W. (1998). Reuniting perception and conception. *Cognition*, 65, 231–262.
- Goldstone, R. L., de Leeuw, J. R., & Landy, D. H. (2015). Fitting perception in and to cognition.

Cognition, 135, 24–29.

- 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, 1124–1142.
- Gray, R., & Cañal-Bruland, R. (in press). Attentional focus, perceived target size, and movement kinematics under performance pressure. *Psychonomic Bulletin & Review*.
- Gray, R., Navia, J. A., & Allsop, J. (2014). Action-specific effects in aviation: What determines judged runway size? *Perception*, 43, 145–154.
- Gregory, R. L. (1980). Perceptions as hypotheses. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, 290, 181–197.
- Hansen, T., Giesel, M., & Gegenfurtner, K. R. (2008). Chromatic discrimination of natural objects. *Journal of Vision*, 8, 2:1–19.
- Hansen, T., Olkkonen, M., Walter, S., & Gegenfurtner, K. R. (2006). Memory modulates color appearance. *Nature Neuroscience*, 9, 1367–1368.
- 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, 1080–1090.
- 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. Rochester, NY, 1925.
- Howard, I. P., & Rogers, B. J. (2002). *Seeing in depth, Volume 2: Depth perception*. Toronto: University of Toronto Press.
- Kanizsa, G. (1985). Seeing and thinking. *Acta Psychologica*, 59, 23–33.
- 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, 431–440.
- 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, 1313–1325.
- Kirsch, W. (2015). Impact of action planning on spatial perception: Attention matters. *Acta Psychologica*, 156, 22–31.
- Klein, G. S., Schlesinger, H. J., & Meister, D. E. (1951). The effect of personal values on perception: An experimental critique. *Psychological Review*, 58, 96–112.
- Klemfuss, N., Prinzmetal, W., & Ivry, R. B. (2012). How does language change perception: A cautionary note. *Frontiers in Psychology*, 3, Article 78.
- Kosslyn, S. M. (2005). Mental images and the brain. *Cognitive Neuropsychology*, 22, 333–347.
- Krauskopf, J., & Karl, G. (1992). Color discrimination and adaptation. *Vision Research*, 32, 2165–2175.
- 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, 978–993.
- 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. *Journal of Neuroscience*, 30, 15254–15261.
- Landis, D., Jones, J. M., & Reiter, J. (1966). Two experiments on perceived size of coins. *Perceptual and Motor Skills*, 23, 719–729.
- Lazarus, R. S., Yousem, H., & Arenberg, D. (1953). Hunger and perception. *Journal of Personality*, 21, 312–328.
- 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*, 90, 11802–11805.
- Lee, D. H., Mirza, R., Flanagan, J. G., & Anderson, A. K. (2014). Optical origins of opposing facial expression actions. *Psychological Science*, 25, 745–752.
- Léger, L., & Chauvet, E. (2015). When canary primes yellow: Effects of semantic memory on overt attention. *Psychonomic Bulletin & Review*, 22, 200–205.

- 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*, 501–512.
- 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 & Psychophysics*, *63*, 676–697.
- Lessard, D. A., Linkenauger, S. A., & Proffitt, D. R. (2009). Look before you leap: Jumping ability affects distance perception. *Perception*, *38*, 1863–1866.
- 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*, 537–547.
- Lo Sciuto, L. A., & Hartley, E. L. (1963). Religious affiliation and open-mindedness in binocular resolution. *Perceptual and Motor Skills*, *17*, 427–430.
- Long, G. M., & Toppino, T. C. (2004). Enduring interest in perceptual ambiguity: Alternating views of reversible figures. *Psychological Bulletin*, *130*, 748–768.
- Lupyan, G. (2012). Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in Psychology*, *3*, Article 54.
- Lupyan, G., & Spivey, M. J. (2008). Perceptual processing is facilitated by ascribing meaning to novel stimuli. *Current Biology*, *18*, R410–R412.
- Lupyan, G., Thompson-Schill, S. L., & Swingley, D. (2010). Conceptual penetration of visual processing. *Psychological Science*, *21*, 682–691.
- Lupyan, G., & Ward, E. J. (2013). Language can boost otherwise unseen objects into visual awareness. *Proceedings of the National Academy of Sciences*, *110*, 14196–14201.
- Machery, E. (in press). Cognitive penetrability: A no-progress report. In J. Zeimbekis & A. Raftopoulos (Eds.), *Cognitive penetrability*. Oxford University Press.
- Macpherson, F. (2012). Cognitive penetration of colour experience: Rethinking the issue in light of an indirect mechanism. *Philosophy and Phenomenological Research*, *84*, 24–62.
- Maner, J. K., Kenrick, D. T., Becker, D. V., Robertson, T. E., Hofer, B., Neuberg, S. L., et al. (2005).

- Functional projection: How fundamental social motives can bias interpersonal perception. *Journal of Personality and Social Psychology*, 88, 63–78.
- Masson, M. E., & Borowsky, R. (1998). More than meets the eye: Context effects in word identification. *Memory & Cognition*, 26, 1245–1269.
- Masters, R., Poolton, J., & van der Kamp, J. (2010). Regard and perceptions of size in soccer: Better is bigger. *Perception*, 39, 1290–1295.
- McCurdy, H. G. (1956). Coin perception studies and the concept of schemata. *Psychological Review*, 63, 160–168.
- 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, 366–376.
- Meteyard, L., Bahrami, B., & Vigliocco, G. (2007). Motion detection and motion verbs language affects low-level visual perception. *Psychological Science*, 18, 1007–1013.
- Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 90, 227–234.
- 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, 1557–1562.
- Most, S. B., Scholl, B. J., Clifford, E. R., & Simons, D. J. (2005). What you see is what you set: Sustained inattention blindness and the capture of awareness. *Psychological Review*, 112, 217–242.
- Nanay, B. (2014). Cognitive penetration and the gallery of indiscernibles. *Frontiers in Psychology*, 5, 1527.
- New, J. J., & German, T. C. (2015). Spiders at the cocktail party: An ancestral threat that surmounts inattention blindness. *Evolution and Human Behavior*, 36, 165–173.
- 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, 935–939.
- 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, 13:1–16.
- Pan, Y., Lin, B., Zhao, Y., & Soto, D. (2014). Working memory biasing of visual perception without awareness. *Attention, Perception, & Psychophysics*, 76, 2051–2062.
- Pascucci, D., & Turatto, M. (2013). Immediate effect of internal reward on visual adaptation. *Psychological Science*, 24, 1317–1322.
- 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, 170–182.
- Peterson, M. A., & Gibson, B. S. (1993). Shape recognition inputs to figure-ground organization in three-dimensional displays. *Cognitive Psychology*, 25, 383–429.
- Peterson, M. A., & Gibson, B. S. (1994). Object recognition contributions to figure-ground organization: operations on outlines and subjective contours. *Perception & Psychophysics*, 56, 551–564.
- 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, 550–557.
- Prinz, J., & Seidel, A. (2012). Alligator or squirrel: Musically induced fear reveals threat in ambiguous figures. *Perception*, 41, 1535–1539.
- Proffitt, D. R. (2006). Embodied perception and the economy of action. *Perspectives on Psychological Science*, 1, 110–122.
- Proffitt, D. R., & Linkenauger, S. A. (2013). Perception viewed as a phenotypic expression. In W. Prinz, M. Beisert, & A. Herwig (Eds.), *Action science: Foundations of an emerging discipline* (pp. 171–198). Cambridge, MA: MIT Press.
- Proffitt, D. R., Stefanucci, J. K., Banton, T., & Epstein, W. (2003). The role of effort in perceiving

- distance. *Psychological Science*, *14*, 106–112.
- Pylyshyn, Z. W. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences*, *22*, 341–365.
- Radel, R., & Clément-Guillotin, C. (2012). Evidence of motivational influences in early visual perception: Hunger modulates conscious access. *Psychological Science*, *23*, 232–234.
- Raftopoulos, A. (2001a). Reentrant neural pathways and the theory-ladenness of perception. *Philosophy of Science*, *68*, S187–S199.
- Raftopoulos, A. (2001b). Is perception informationally encapsulated? The issue of the theory-ladenness of perception. *Cognitive Science*, *25*, 423–451.
- Raftopoulos, A., & Zeimbekis, J. (Eds.). (in press). *Cognitive penetrability*. Cambridge, MA: MIT Press.
- Ramachandran, V. (1988). Perception of shape from shading. *Nature*, *331*, 163–166.
- 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*, 401–414.
- Riener, C. R., Stefanucci, J. K., Proffitt, D. R., & Clore, G. L. (2011). An effect of mood on the perception of geographical slant. *Cognition & Emotion*, *25*, 174–182.
- 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*, 171–178.
- Rock, I. (1983). *The logic of perception*. Cambridge, MA: MIT Press.
- Rolfs, M., Dambacher, M., & Cavanagh, P. (2013). Visual adaptation of the perception of causality. *Current Biology*, *23*, R202–R204.
- Rolls, E. T. (2008). Top-down control of visual perception: Attention in natural vision. *Perception*, *37*, 333–354.
- Rosenthal, R., & Rubin, D. B. (1978). Interpersonal expectancy effects: The first 345 studies. *Behavioral and Brain Sciences*, *1*, 377–386.
- Rust, N. C., Mante, V., Simoncelli, E. P., & Movshon, J. A. (2006). How MT cells analyze the

- motion of visual patterns. *Nature Neuroscience*, 9, 1421–1431.
- 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, 464–482.
- Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, 80, 1–46.
- Scholl, B. J., & Gao, T. (2013). Perceiving animacy and intentionality: Visual processing or higher-level judgment? In M. D. Rutherford & V. A. Kuhlmeier (Eds.), *Social Perception: Detection and Interpretation of Animacy, Agency, and Intention* (pp. 197–230). Cambridge, MA: MIT Press.
- Scholl, B. J., & Leslie, A. M. (1999). Modularity, development and ‘theory of mind.’ *Mind and Language*, 14, 131–153.
- Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, 4, 299–309.
- Sekuler, R., Sekuler, A. B., & Lau, R. (1997). Sound alters visual motion perception. *Nature*, 385, 308–308.
- Shaffer, D. M., & Flint, M. (2011). Escalating slant: Increasing physiological potential does not reduce slant overestimates. *Psychological Science*, 22, 209–211.
- 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, 147–155.
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature*, 408, 788.
- Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, 14, 147–152.
- Siegel, S. (2012). Cognitive penetrability and perceptual justification. *Noûs*, 46, 201–222.
- Simons, D. J. (1996). In sight, out of mind: When object representations fail. *Psychological Science*, 7, 301–305.
- Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentive blindness for dynamic events. *Perception*, 28, 1059–1074.

- 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*, *109*, 19614–19619.
- 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*, 450–452.
- 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.
- 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.
- Stefanucci, J. K., & Geuss, M. N. (2009). Big people, little world: The body influences size perception. *Perception*, *38*, 1782–1795.
- 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*, 424–438.
- Stefanucci, J. K., & Storbeck, J. J. (2009). Don't look down: Emotional arousal elevates height perception. *Journal of Experimental Psychology: General*, *138*, 131–145.
- Stokes, D. (2014). Cognitive penetration and the perception of art. *Dialectica*, *68*, 1–34.
- Storbeck, J., & Stefanucci, J. K. (2014). Conditions under which arousal does and does not elevate height estimates. *PLoS One*, *9*, e92024.
- Sugovic, M., & Witt, J. K. (2013). An older view on distance perception: Older adults perceive walkable extents as farther. *Experimental Brain Research*, *226*, 383–391.
- Taylor, J. E. T., Witt, J. K., & Sugovic, M. (2011). When walls are no longer barriers: Perception of wall height in parkour. *Perception*, *40*, 757–760.
- 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.
- Toppino, T. C. (2003). Reversible-figure perception: Mechanisms of intentional control. *Perception & Psychophysics*, *65*, 1285–1295.
- Tse, P. U. (2005). Voluntary attention modulates the brightness of overlapping transparent

- surfaces. *Vision Research*, 45, 1095–1098.
- Tse, P. U., & Cavanagh, P. (2000). Chinese and Americans see opposite apparent motions in a Chinese character. *Cognition*, 74, B27–B32.
- 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, 293–299.
- 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, 304–310.
- Vetter, P., & Newen, A. (2014). Varieties of cognitive penetration in visual perception. *Consciousness and Cognition*, 27, 62–75.
- Vickery, T. J., Chun, M. M., & Lee, D. (2011). Ubiquity and specificity of reinforcement signals throughout the human brain. *Neuron*, 72, 166–177.
- Wakslak, C. J., & Kim, B. K. (in press). Controllable objects seem closer. *Journal of Experimental Psychology: General*.
- Ward, E. J., & Scholl, B. J. (2015). Inattention blindness reflects limitations on perception, not memory: Evidence from repeated failures of awareness. *Psychonomic Bulletin & Review*, 22, 722–727.
- Webster, M. A., & Kay, P. (2012). Color categories and color appearance. *Cognition*, 122, 375–392.
- Wesp, R., Cichello, P., Gracia, E. B., & Davis, K. (2004). Observing and engaging in purposeful actions with objects influences estimates of their size. *Perception & Psychophysics*, 66, 1261–1267.
- Wesp, R., & Gasper, J. (2012). Is size misperception of targets simply justification for poor performance? *Perception*, 41, 994–996.
- Witt, J. K. (2011a). Action's effect on perception. *Current Directions in Psychological Science*, 20, 201–206.
- 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, 1148–1156.
- Witt, J. K., & Dorsch, T. E. (2009). Kicking to bigger uprights: Field goal kicking performance influences perceived size. *Perception*, 38, 1328–1340.
- 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 & Review*, 15, 581–585.
- Witt, J. K., & Proffitt, D. R. (2005). See the ball, hit the ball. *Psychological Science*, 16, 937–938.
- Witt, J. K., Proffitt, D. R., & Epstein, W. (2004). Perceiving distance: A role of effort and intent. *Perception*, 33, 577–590.
- 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, 880–888.
- 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, 1153–1160.
- Witt, J. K., Schuck, D. M., & Taylor, J. E. T. (2011). Action-specific effects underwater. *Perception*, 40, 530–537.
- Witt, J. K., & Sugovic, M. (2010). Performance and ease influence perceived speed. *Perception*, 39, 1341–1353.
- Witzel, C., Valkova, H., Hansen, T., & Gegenfurtner, K. R. (2011). Object knowledge modulates colour appearance. *i-Perception*, 2, 13–49.
- 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, 1104–1117.
- Yee, E., Ahmed, S. Z., & Thompson-Schill, S. L. (2012). Colorless green ideas (can) prime furiously. *Psychological Science*, 23, 364–369.
- Zadra, J. R., & Clore, G. L. (2011). Emotion and perception: The role of affective information. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2, 676–685.

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, 660–665.

Author Note

For helpful conversation and/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.

Figure Captions

Figure 1. Examples of lightness illusions can be subjectively appreciated as “demonstrations” (for references and explanations, see Adelson, 2000). (A) The two columns of grey rectangles have the same objective 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.

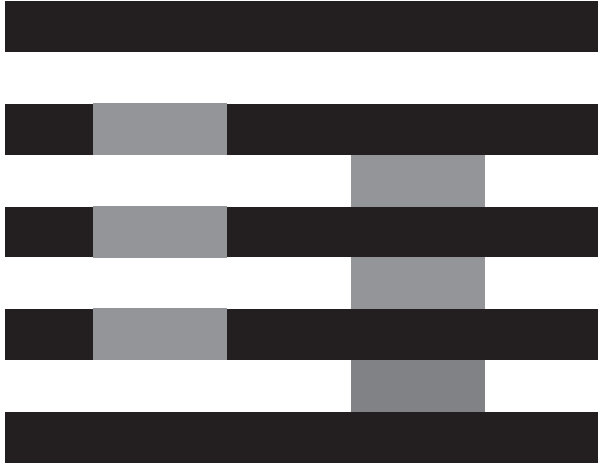
Figure 2. Diagrams or depictions of various possible top-down effects on perception: (A) Figure-ground assignment is biased toward familiar shapes, such as the profile of a woman (Peterson & Gibson, 1993). (B) Being thirsty (as a result of eating salty pretzels) makes ambiguous surfaces look more transparent (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 (Wesp & Gasper, 2004; Cañal-Bruland et al., 2010). (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 look to have their category’s 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 & Gegenfurtner, 2006). (L) Conceptual similarity enhances size-contrast illusions (Coren & Enns, 1993). (M) Labeling certain ‘squiggle’ 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 6th 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 bean-bag 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).

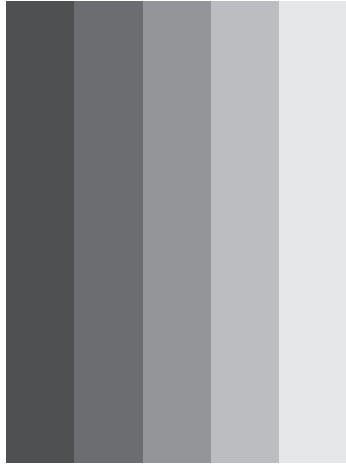
Figure 3. (A) The face stimuli (matched in mean luminance) from Experiment 1 of Levin and Banaji (2006). (B) The blurred versions of the same stimuli used in Firestone & Scholl (2015a), which preserved the match in mean luminance, while obscuring the race information.

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

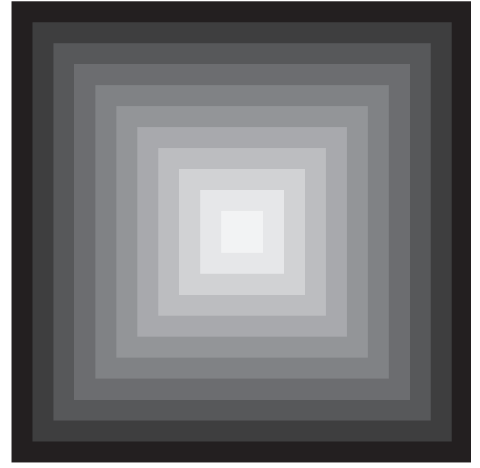
A



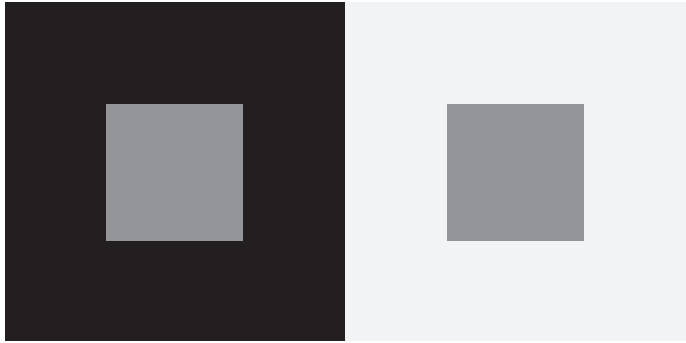
B



C



D



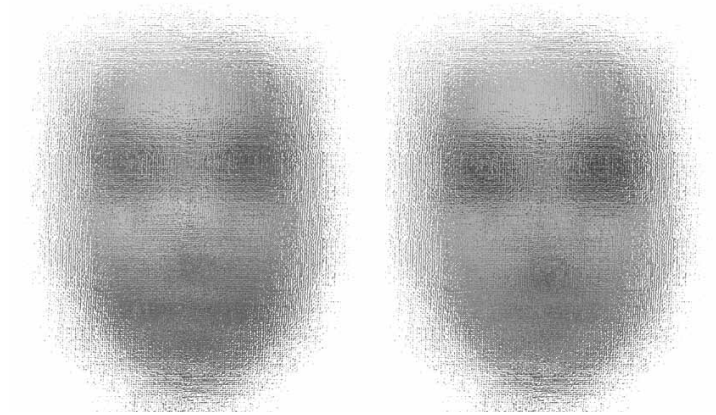
E



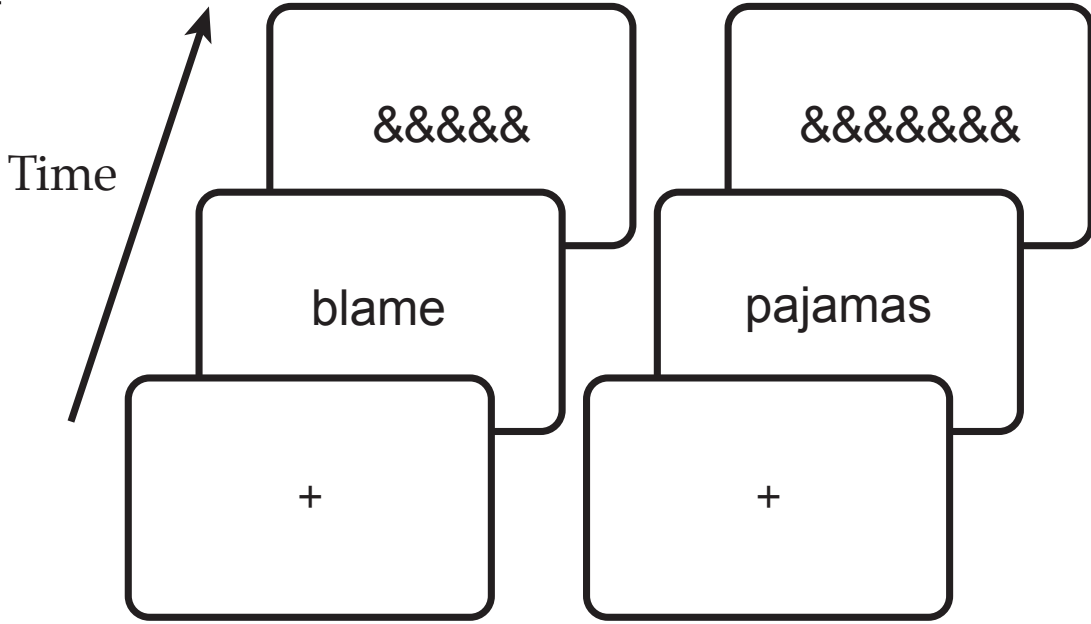
A



B



A



B

