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

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

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

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

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

ACKNOWLEDGMENT

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

Perception, as you make it

doi:10.1017/S0140525X15002678, e260

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

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

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

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

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

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

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

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

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

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

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

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

An action-specific effect on perception that avoids all pitfalls

doi:10.1017/S0140525X15002563, e261

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

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

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

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

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

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

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

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

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