Perception, As You Make It

Behavioral & Brain Sciences commentary on Chaz Firestone & Brian Scholl, "Cognition does not affect perception: Evaluating the evidence for 'top-down' effects"

David W. Vinson, University of California, Merced,
Drew H. Abney, University of California, Merced
Dima Amso, Brown University
Anthony Chemero, University of Cincinnati
James E. Cutting, Cornell University
Rick Dale, University of California, Merced
Jonathan B. Freeman, New York University
Laurie B. Feldman, The University at Albany, SUNY
Karl J. Friston, University College London
Shaun Gallagher, University of Memphis
J. Scott Jordan, Illinois State University
Liad Mudrik, Tel Aviv University
Sasha Ondobaka, University College London
Daniel C. Richardson, University College London
Ladan Shams, University of California, Los Angeles
Maggie Shiffrar, California State University, Northridge
Michael J. Spivey, University of California, Merced

Abstract: The main question F&S pose is whether "what and how we see is functionally independent from what and how we think, know, desire, act, etc." We synthesize a collection of concerns from an interdisciplinary set of co-authors 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 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, and 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, one can cognitively control which of these you see. Admittedly, cognition cannot "penetrate" our perception to turn straight lines into curved ones; and clearly we cannot see a young woman in the Jastrow 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, Pardo, Piorkowski & Rubin, 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 the visual figures above, this is not a bistable phenomenon; once heard as speech one cannot fully go back and hear these signals as the squeaks one heard before. Is this not top-down?

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

Secondly, visual neuroscience shows numerous examples of feedback projections to visual cortex, and feedback influences on visual neural processing ignored by F&S. 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 (Clavignier, Falchier, & Kennedy, 2004) and the amygdala (Amaral, Behniea, & Kelly, 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 (p. 9).

Neurons in V1 exhibit differential responses to the same visual input under a variety of contextual modulations (e.g., David, Vinje, & Gallant, 2004; Hupé et al. 1998; Kapadia, Ito, Gilbert, & Westheimer, 1995; Motter, 1993). Numerous studies with adults have established that selective attention enhances processing of information at the attended location, and suppresses distraction (Gandhi, Heeger, & Boynton, 1999; Kastner, Pinsk, De Weerd, Desimone, & Ungerleider, 1999; Markant, Worden, & Amso, 2015; Slotnick, Schwarzbach, & Yantis, 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, Clapp, Chadick, & Gazzaley, 2010; Uncapher & Rugg, 2009; Zanto & Gazzaley, 2009) and infants (Markant & Amso, 2013, 2015; Markant, Oakes & Amso, 2015). Recent data indicate that attentional biases can function at higher levels in the cognitive hierarchy (Chua & Gautier, 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, Hinton, Neal & Zemel, 1995; Friston, 2010; Gregory, 1980; Jordan, 2013; Kastner & Ungerleider, 2001; Kveraga, Ghuman, & Bar, 2007; 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." Indeed, F&S's argument is heavily recycled from a previous BBS contribution (Pylyshyn, 1999). Despite their attempt to distinguish their contribution from this one, it suffers from very similar weaknesses identified by past commentary (e.g., Bruce, Langton, & Hill, 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." Especially in the case of visual perception, that is exactly what has been happening in the field for these past few decades.

References

Amaral, D. G., Behniea, H., & Kelly, J. L. (2003). Topographic organization of projections from the amygdala to the visual cortex in the macaque monkey. *Neuroscience, 118*(4), 1099-1120.

Anton-Erxleben, K., & Carrasco, M. (2013). Attentional enhancement of spatial resolution: linking behavioural and neurophysiological evidence. *Nature Reviews Neuroscience*, *14*(3), 188-200.

Bruce, V., Langton, S., & Hill, H. (1999). Complexities of face perception and categorisation. *Behavioral and Brain Sciences, 22*(3), 369-370.

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

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, *51*(13), 1484-1525.

Clavagnier, S., Falchier, A., & Kennedy, H. (2004). Long-distance feedback projections to area V1: implications for multisensory integration, spatial awareness, and visual consciousness. *Cognitive, Affective, & Behavioral Neuroscience, 4*(2), 117-126.

Cavanagh, P. (1999). The cognitive penetrability of cognition. *Behavioral and Brain Sciences, 22*(3), 370-371.

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

David, S. V., Vinje, W. E., & Gallant, J. L. (2004). Natural stimulus statistics alter the receptive field structure of v1 neurons. *The Journal of Neuroscience, 24*(31), 6991-7006.

Darwin, C. J. (1997). Auditory grouping. *Trends in Cognitive Sciences*, *1*(9), 327-333.

Dayan, Hinton G. E., Neal R., Zemel R. (1995). The Helmholtz machine. *Neural Computation, 7*(5), 889–904.

Friston K. J. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience, 11*(2), 127–138.

Fodor, J. A. (1983). The modularity of mind: An essay on faculty psychology. MIT press.

Gandhi, S. P., Heeger, D. J., & Boynton, G. M. (1999). Spatial attention affects brain activity in human primary visual cortex. *Proceedings of the National Academy of Sciences, 96*(6), 3314-3319.

Gregory, R. L. (1980). Perceptions as hypotheses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *290*(1038), 181-197.

Hill, W. E. (1915). My wife and my mother-in-law. *Puck*, *6*.

Hupé, J. M., James, A. C., Payne, B. R., Lomber, S. G., Girard, P., & Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature, 394*(6695), 784-787.

Jastrow, J. (1899). The mind's eye. *Popular Science Monthly, 54*, 299-312.

Jordan, J. S. (2013). The wild ways of conscious will: what we do, how we do it, and why it has meaning. *Frontiers in Psychology*, *4*, 574.

Kapadia, M. K., Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys. *Neuron, 15*(4), 843-856.

Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron, 22*(4), 751-761.

Kastner, S., & Ungerleider, L. G. (2001). The neural basis of biased competition in human visual cortex. *Neuropsychologia, 39*(12), 1263-1276.

Kveraga, K., Ghuman, A. S., & Bar, M. (2007). Top-down predictions in the cognitive brain. *Brain and Cognition, 65*(2), 145-168.

Lupyan, G., & Spivey, M. J. (2010). Making the invisible visible: Verbal but not visual cues enhance visual detection. *PloS One, 5*(7), e11452.

Markant, J., & Amso, D. (2013). Selective memories: infants' encoding is enhanced in selection via suppression. *Developmental Science, 16*(6), 926-940.

Markant, J. & Amso, D. (2015). The development of selective attention orienting is an agent of change in learning and memory efficacy. *Infancy*, 1-23. doi: 10.1111/infa.12100

Markant, J., Worden, M. S., & Amso, D. (2015). Not all attention orienting is created equal: Recognition memory is enhanced when attention orienting involves distractor suppression. *Neurobiology of Learning and Memory, 120*, 28-40.

Markant, J., Oakes, L. M., & Amso, D. (2015). Visual selective attention biases contribute to the other-race effect among 9-month-old infants. *Developmental Psychobiology*. doi: 10.1002/dev.21375

Motter, B. C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *Journal of Neurophysiology, 70*(3), 909-919.

Spratling, M. W. (2010). Predictive coding as a model of response properties in cortical area V1. *The Journal of Neuroscience, 30*(9), 3531-3543.

Pylyshyn, Z. (1999). Is vision continuous with cognition?: The case for cognitive impenetrability of visual perception. *Behavioral and brain sciences*, *22*(3), 341-365.

Rao R. P., & Ballard D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience 2*(1), 79–87.

Remez, R. E., Pardo, J. S., Piorkowski, R. L., & Rubin, P. E. (2001). On the bistability of sine wave analogues of speech. *Psychological Science*, *12*(1), 24-29.

Roepstorff, A., & Frith, C. (2004). What’s at the top in the top-down control of action? Script-sharing and ‘top-top’ control of action in cognitive experiments. *Psychological Research, 68*(2-3), 189-198.

Rutman, A. M., Clapp, W. C., Chadick, J. Z., & Gazzaley, A. (2010). Early top–down control of visual processing predicts working memory performance. *Journal of Cognitive Neuroscience, 22*(6), 1224-1234.

Slotnick, S. D., Schwarzbach, J., & Yantis, S. (2003). Attentional inhibition of visual processing in human striate and extrastriate cortex. *Neuroimage, 19*(4), 1602-1611.

Uncapher, M. R., & Rugg, M. D. (2009). Selecting for memory? The influence of selective attention on the mnemonic binding of contextual information. *The Journal of Neuroscience, 29*(25), 8270-8279.

Zanto, T. P., & Gazzaley, A. (2009). Neural suppression of irrelevant information underlies optimal working memory performance. *The Journal of Neuroscience, 29*(10), 3059-3066.

Zhang, P., Jamison, K., Engel, S., He, B., & He, S. (2011). Binocular rivalry requires visual attention. *Neuron, 71*(2), 362-369.