

Comment on BBS target article by **Firestone & Scholl**

**Word count**

Abstract (max. 60):	59
Main text (max. 1k):	1000
References:	844
Entire text:	1992

# Attention and multisensory modulation argue against total encapsulation

<sup>1,3</sup>Benjamin de Haas, <sup>1,3</sup>Dietrich Samuel Schwarzkopf, <sup>1,2</sup>Geraint Rees

<sup>1</sup>Institute of Cognitive Neuroscience, 17 Queen Square, WC1N 3AR, <sup>2</sup>Wellcome Trust Centre for Neuroimaging, 12 Queen Square, WC1N 3BG, <sup>3</sup>Experimental Psychology, 26 Bedford Way, WC1H 0AP

University College London, UK

Telephone: +44-20-7679-5349

E-mail:

[benjamin.haas.09@ucl.ac.uk](mailto:benjamin.haas.09@ucl.ac.uk)  
[s.schwarzkopf@ucl.ac.uk](mailto:s.schwarzkopf@ucl.ac.uk)  
[g.rees@ucl.ac.uk](mailto:g.rees@ucl.ac.uk)

Web:

<https://www.ucl.ac.uk/pals/research/experimental-psychology/person/benjamin-de-haas/>  
<https://sampendu.wordpress.com/sam-schwarzkopf/>  
<http://www.fil.ion.ucl.ac.uk/~grees/>

## Abstract

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

## Main Text

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

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

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

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

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

## **Acknowledgments**

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

## **References**

- Abrams, J., Barbot, A., & Carrasco, M. (2010). Voluntary attention increases perceived spatial frequency. *Attention, Perception & Psychophysics*, 72(6), 1510–21. doi:10.3758/APP.72.6.1510
- Anton-Erxleben, K., Abrams, J., & Carrasco, M. (2011). Equality judgments cannot distinguish between attention effects on appearance and criterion: a reply to Schneider (2011). *Journal of Vision*, 11(13), 8. doi:10.1167/11.13.8
- Anton-Erxleben, K., & Carrasco, M. (2013). Attentional enhancement of spatial resolution: linking behavioural and neurophysiological evidence. *Nature Reviews. Neuroscience*, 14(3), 188–200. doi:10.1038/nrn3443
- Baruch, O., & Yeshurun, Y. (2014). Attentional attraction of receptive fields can explain spatial and temporal effects of attention. *Visual Cognition*, 22(5), 704–736. doi:10.1080/13506285.2014.911235
- Berger, T. D., Martelli, M., & Pelli, D. G. (2003). Flicker flutter: is an illusory event as good as the real thing? *Journal of Vision*, 3(6), 406–12. doi:10.1167/3.6.1
- Bruns, P., Maiworm, M., & Röder, B. (2014). Reward expectation influences audiovisual spatial integration. *Attention, Perception & Psychophysics*, 76(6), 1815–27. doi:10.3758/s13414-014-0699-y
- Carmel, D., Thorne, J. D., Rees, G., & Lavie, N. (2011). Perceptual load alters visual excitability. *Journal of Experimental Psychology. Human Perception and Performance*, 37(5), 1350–60. doi:10.1037/a0024320
- Carrasco, M., Fuller, S., & Ling, S. (2008). Transient attention does increase perceived contrast of suprathreshold stimuli: A reply to Prinzmetal, Long, and Leonhardt (2008). *Perception & Psychophysics*, 70(7), 1151–1164. doi:10.3758/PP.70.7.1151
- de Haas, B., Cecere, R., Cullen, H., Driver, J., & Romei, V. (2013). The duration of a co-occurring sound modulates visual detection performance in humans. *PloS One*, 8(1), e54789. doi:10.1371/journal.pone.0054789
- de Haas, B., Kanai, R., Jalkanen, L., & Rees, G. (2012). Grey matter volume in early human visual cortex predicts proneness to the sound-induced flash illusion. *Proceedings. Biological Sciences / The Royal Society*, 279(1749), 4955–61. doi:10.1098/rspb.2012.2132
- de Haas, B., & Rees, G. (2010). Multiple stages of cross-modal integration in visual processing: Comment on “Crossmodal influences on visual perception” by L. Shams & R. Kim. *Physics of Life Reviews*.
- de Haas, B., Schwarzkopf, D. S., Anderson, E. J., & Rees, G. (2014). Perceptual load affects spatial tuning of neuronal populations in human early visual cortex.

*Current Biology : CB*, 24(2), R66–7. doi:10.1016/j.cub.2013.11.061

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

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

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

Dumoulin, S. O., & Wandell, B. A. (2008). Population receptive field estimates in human visual cortex. *NeuroImage*, 39(2), 647–60. doi:10.1016/j.neuroimage.2007.09.034

Fodor, J. A. (1983). *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press.

Gau, R., & Noppeney, U. (2015). How prior expectations shape multisensory perception. *NeuroImage*. doi:10.1016/j.neuroimage.2015.09.045

Klein, B. P., Harvey, B. M., & Dumoulin, S. O. (2014). Attraction of position preference by spatial attention throughout human visual cortex. *Neuron*, 84(1), 227–37. doi:10.1016/j.neuron.2014.08.047

Lavie, N. (2005). Distracted and confused?: selective attention under load. *Trends in Cognitive Sciences*, 9(2), 75–82. doi:10.1016/j.tics.2004.12.004

Liu, T., Abrams, J., & Carrasco, M. (2009). Voluntary attention enhances contrast appearance. *Psychological Science*.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588), 746–8.

Montagna, B., Pestilli, F., & Carrasco, M. (2009). Attention trades off spatial acuity. *Vision Research*, 49(7), 735–45.

Navarra, J., Alsius, A., Soto-Faraco, S., & Spence, C. (2010). Assessing the role of attention in the audiovisual integration of speech. *Information Fusion*, 11(1), 4–11. doi:10.1016/j.inffus.2009.04.001

Pilgramm, S., de Haas, B., Helm, F., Zentgraf, K., Stark, R., Munzert, J., & Krüger, B. (2015). Motor imagery of hand actions: Decoding the content of motor imagery from brain activity in frontal and parietal motor areas. *Human Brain Mapping*. doi:10.1002/hbm.23015

- Rees, G., Frith, C. D., & Lavie, N. (1997). Modulating irrelevant motion perception by varying attentional load in an unrelated task. *Science (New York, N.Y.)*, 278(5343), 1616–9.
- Romei, V., de Haas, B., Mok, R. M., & Driver, J. (2011). Auditory Stimulus Timing Influences Perceived duration of Co-Occurring Visual Stimuli. *Frontiers in Psychology*, 2, 215. doi:10.3389/fpsyg.2011.00215
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions: What you see is what you hear. *Nature*, 408(December), 788.
- Suzuki, S., & Cavanagh, P. (1997). Focused attention distorts visual space: an attentional repulsion effect. *Journal of Experimental Psychology. Human Perception and Performance*, 23(2), 443–63.
- Thomas, G. (1941). Experimental study of the influence of vision on sound localization. *Journal of Experimental Psychology*.
- Vetter, P., Smith, F. W., & Muckli, L. (2014). Decoding sound and imagery content in early visual cortex. *Current Biology : CB*, 24(11), 1256–62. doi:10.1016/j.cub.2014.04.020
- Watkins, S., Shams, L., Tanaka, S., Haynes, J.-D., & Rees, G. (2006). Sound alters activity in human V1 in association with illusory visual perception. *NeuroImage*, 31(3), 1247–56. doi:10.1016/j.neuroimage.2006.01.016
- Womelsdorf, T., Anton-Erxleben, K., & Treue, S. (2008). Receptive field shift and shrinkage in macaque middle temporal area through attentional gain modulation. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 28(36), 8934–44. doi:10.1523/JNEUROSCI.4030-07.2008