

## Journal Club

**Editor's Note:** These short, critical reviews of recent papers in the *Journal*, written exclusively by graduate students or postdoctoral fellows, are intended to summarize the important findings of the paper and provide additional insight and commentary. For more information on the format and purpose of the Journal Club, please see [http://www.jneurosci.org/misc/ifa\\_features.shtml](http://www.jneurosci.org/misc/ifa_features.shtml).

## Predictive Coding Explains Auditory and Tactile Influences on Vision during Binocular Rivalry

Bradley N. Jack and Graeme Hacker

Discipline of Psychology, School of Health and Human Sciences, Southern Cross University, Coffs Harbour, Australia  
Review of Lunghi et al.

When the two eyes are presented with dissimilar images, such as a picture of a face to the left eye and a picture of a house to the right eye, something remarkable happens. Observers report seeing the picture of the face for a few seconds, then the picture of the house, then the face again, and so on, for as long as he or she cares to look. This phenomenon is called binocular rivalry (Alais and Blake, 2004; Miller, 2013), and although the neurocognitive mechanism that explains this visual effect remains unsolved, there is a trend toward approaches that combine top-down and bottom-up processes in the brain (Blake and Logothetis, 2002; Tong et al., 2006).

Despite there being a rich literature on the topic of binocular rivalry, only a handful of studies have investigated the influences of multisensory stimuli on vision during binocular rivalry (Zhou et al., 2010). This is surprising, because perception is almost always multisensory. That is, a stable representation of the environment often consists of a combination of visual, auditory, tactile, gustatory, and/or olfactory inputs. Recently, a study in *The Journal of Neuroscience* sought to investigate the influences of multisensory stimuli on vision during binocular rivalry. In particular, Lunghi et al. (2014) set out to understand whether auditory, tactile, and

auditory/tactile stimuli could influence percept selection and percept alternation during binocular rivalry.

To instigate binocular rivalry, Lunghi et al. (2014) presented spatially identical, contrast-modulating, random-noise patterns to the two eyes. The stimuli differed only in the temporal frequency in which they modulated in contrast, one at 3.75 Hz and the other at 15 Hz. This type of binocular-rivalry stimulus is known to produce alternations in vision that are comparable to alternations caused by spatially dissimilar stimuli, such as those of a face and a house (Alais and Parker, 2012). During dichoptic viewing of the stimuli, Lunghi et al. (2014) presented auditory, tactile, or auditory/tactile stimuli intermittently. Their auditory and tactile stimuli consisted of amplitude-modulated sine waves at either 3.75 or 15 Hz—these frequencies are the same as the ones used by the visual stimuli, and were presented through headphones (auditory stimuli) or as vibrations (tactile stimuli). Finally, Lunghi et al. (2014) manipulated the intensity of the auditory and tactile stimuli, resulting in modulation depths of 33%, 66%, and 100%. They did this to assess the effects of auditory and tactile stimulus strength on vision during binocular rivalry.

Lunghi et al. (2014) found that presenting auditory and tactile inputs that were congruent with an observer's percept, as reported by key presses, during binocular rivalry resulted in an increased probability of maintaining that percept, compared with when there was no audi-

tory or tactile input, whereas inputs that were incongruent resulted in an increased probability of switching that percept to the congruent alternative. Furthermore, the probability of percept switching or percept maintaining was influenced by auditory and tactile stimulus strength: inputs with a modulation depth of 100% resulted in a higher probability of percept switching ( $\sim 0.67$ ) or percept maintaining ( $\sim 0.4$ ) than inputs with a modulation depth of 33% ( $\sim 0.5$  for percept switching and  $\sim 0.3$  for percept maintaining). Lunghi et al. (2014) also found that when auditory and tactile inputs were presented simultaneously at a modulation depth of 33% each, the probability of percept switching or percept maintaining was comparable to the results obtained for auditory only or tactile only inputs with a modulation depth of 100%. Finally, when auditory and tactile inputs were presented simultaneously at a modulation depth of 100% each but in anti-phase (i.e., the auditory input was congruent with one visual stimulus whereas the tactile input was congruent with the other), there was no effect on the probability of percept switching or percept maintaining (for more details, see Lunghi et al., 2014, their Fig. 1). Lunghi et al. (2014) concluded that there must be a common, bottom-up, neural mechanism for the integration of visual, auditory, and tactile inputs, and that this mechanism is used to deliver a stable representation of the environment.

We agree that there must be a common mechanism for the integration of multi-

Received March 15, 2014; revised April 1, 2014; accepted April 2, 2014.

Correspondence should be addressed to Bradley N. Jack, Discipline of Psychology, Southern Cross University, Hogbin Drive, Coffs Harbour NSW 2450, Australia. E-mail: [bradley.jack@scu.edu.au](mailto:bradley.jack@scu.edu.au).

DOI:10.1523/JNEUROSCI.1040-14.2014

Copyright © 2014 the authors 0270-6474/14/346423-02\$15.00/0

sensory inputs, but we suspect this mechanism alone would be insufficient to govern visual perception, because it does not account for top-down influences on vision during binocular rivalry (Blake and Logothetis, 2002; Tong et al., 2006). Below, we use predictive coding to provide an epistemological explanation for why, when presented with auditory, tactile, or auditory/tactile inputs during binocular rivalry, visual awareness or one of the other visual input is maintained or switched.

Predictive coding is a framework for understanding how the brain delivers a stable representation of the environment. Theories of predictive coding advocate that the cognitive system is structured in a hierarchical manner, that all perception is a testable hypothesis, and that this hypothesis is a compromise between top-down predictions and bottom-up sensory input. According to predictive coding, the brain delivers a stable representation of the environment by using top-down predictions to reduce redundant processing of repetitive stimuli from bottom-up sensory input and by processing only what is not predicted—prediction-error (Rao and Ballard, 1999; Friston, 2005). Indeed, the concepts of predictive coding have also been used to explain binocular rivalry (Hohwy et al., 2008).

Hohwy et al. (2008) argue that the selection problem in binocular rivalry, why one stimulus is selected for perception rather than the other or both, occurs because one of the images (i.e., a face) has a higher prior, or likelihood, than the other image (i.e., a house) or than both images superimposed (i.e., a face-house). In addition, Hohwy et al. (2008) solve the alternation problem, why the images alternate in visual perception over time, using prediction-error. If one of the images accounts for approximately half of the total visual input, then the other half of visual input is unexplained and is therefore prediction-error. According to theories of predictive coding, prediction-error is passed up the hierarchy until it is resolved by higher-level neurons. When this happens, there is a change in visual perception, such that the image that was previously suppressed from awareness is now perceived and the image that was previously perceived is now suppressed. Furthermore, the image that is now suppressed

is the new source of prediction-error in sensory input. This pattern repeats itself for as long as one cares to look at binocular-rivalry stimuli. We think this explanation of binocular rivalry could be extended to account for the findings reported by Lunghi et al. (2014).

Lunghi et al. (2014) found that presenting auditory and tactile inputs that were congruent with an observer's percept during binocular rivalry resulted in an increased probability of maintaining that percept. This is consistent with predictive coding, because more than half of the total sensory input (visual and auditory or tactile vs visual only) is congruent with the brain's hypothesis about the visual environment; thus, the current predictive model is a good representation of sensory input, and is maintained. Of course, this percept will not be maintained forever, because the brain needs to eventually account for prediction-error from the suppressed visual stimulus. Lunghi et al. (2014) also found that presenting auditory and tactile inputs that were incongruent with an observer's percept resulted in an increased probability of switching that percept. Again, this is consistent with predictive coding, because more than half of the total sensory input is prediction-error; thus, the current predictive model is a poor representation of sensory input, and it is replaced by another model that better accounts for sensory input, forcing a change in visual perception.

Predictive coding can also explain why the probability of percept switching or percept maintaining was influenced by auditory and tactile stimulus strength. If percept maintaining and percept switching are the result of decreased and increased prediction-error, respectively, then reducing the modulation depth of auditory and tactile inputs from 100% to 33% decreases the amount of total sensory input that the current predictive model accounts for. Thus, one would expect a lower probability of percept switching or percept maintaining for auditory and tactile inputs at a lower modulation depth than at a higher modulation depth. This reasoning can also be used to explain why auditory and tactile inputs presented simultaneously at a modulation depth of 33% each yield results that are much higher than those obtained for auditory-only or tactile-only inputs with a modula-

tion depth of 33%, as well as explain why auditory and tactile inputs presented simultaneously at a modulation depth of 100% each but in anti-phase have no effect on vision.

Of course, the concept of predictive coding do not detract from Lunghi et al.'s (2014) conclusion: that perception of one or the other stimulus during binocular rivalry is driven by a common, bottom-up, multisensory mechanism. Rather, predictive coding complements Lunghi et al.'s (2014) bottom-up explanation by acknowledging top-down influences on vision during binocular rivalry. Indeed, any explanation of percept selection and percept alternation during binocular rivalry is likely to be incomplete without including top-down mechanisms of perception (Blake and Logothetis, 2002; Tong et al., 2006). In conclusion, predictive coding is a useful framework for understanding the interaction between different levels of the visual hierarchy that allow our brains to deliver a stable representation of the environment.

## References

- Alais D, Parker A (2012) Binocular rivalry produced by temporal frequency differences. *Front Hum Neurosci* 6:227. [CrossRef Medline](#)
- Alais D, Blake R (2004) Binocular rivalry. Cambridge, MA: MIT.
- Blake R, Logothetis N (2002) Visual competition. *Nat Rev Neurosci* 3:13–21. [CrossRef Medline](#)
- Friston K (2005) A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci* 360:815–836. [CrossRef Medline](#)
- Hohwy J, Roepstorff A, Friston K (2008) Predictive coding explains binocular rivalry: an epistemological review. *Cognition* 108:687–701. [CrossRef Medline](#)
- Lunghi C, Morrone MC, Alais D (2014) Auditory and tactile signals combine to influence vision during binocular rivalry. *J Neurosci* 34:784–792. [CrossRef Medline](#)
- Miller SM (2013) The constitution of visual consciousness: Lessons from binocular rivalry. Amsterdam: John Benjamin's Publishing.
- Rao RP, Ballard DH (1999) Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci* 2:79–87. [Medline](#)
- Tong F, Meng M, Blake R (2006) Neural bases of binocular rivalry. *Trends in Cognitive Sciences* 10:502–511. [CrossRef Medline](#)
- Zhou W, Jiang Y, He S, Chen D (2010) Olfaction modulates visual perception in binocular rivalry. *Curr Biol* 20:1356–1358. [CrossRef Medline](#)