

1 Introduction

How does the brain predict what happens from one moment to the next? This is an important question in research on perception across all modalities that is surprisingly often overlooked in the fields of psychology and neuroscience. For example, most experiments are designed to measure evoked responses to a randomly chosen, isolated stimulus under the tacit assumption that response variability is irrelevant noise that should be averaged out across many presentations. Computational models of perceptual processing often operate under similar assumptions in which stimuli are presented as random “snapshots” from which some common set of features should be learned to minimize representational variability across presentations (e.g., Serre, Oliva, & Poggio, 2007; Mutch & Lowe, 2008; although see Foldiak, 1991, for a notable exception). These experimental and computational assumptions stand in contrast to the event structure of the physical world, which is deterministic from one moment to the next.

An equally important question is concerned with how the brain knows when to make its next prediction. Predicting what happens next requires integrating information over some time frame and using the result to drive the actual prediction, but when should integration start? And for how long? Perceptual processing has been shown to undergo temporal fluctuations and in extreme cases, stimuli can be rendered imperceptible if presented during one of these fluctuations (Busch, Dubois, & VanRullen, 2009; VanRullen, Busch, Drewes, & Dubois, 2011; Mathewson et al., 2011). Again, many laboratory experiments tend not to be concerned with these temporal fluctuations, as they simply add variability to responses that will average out over a large number of trials and can be mitigated by design decisions such as using a fixation cross to denote the start of a trial. Computational models, similarly, often completely ignore time altogether, although recent advances in spiking models of perceptual processing (e.g., Masquelier & Thorpe, 2007) are beginning to address this issue.

The goal of the proposed work is to begin systematically investigating the neural mechanisms and circuitry related to prediction and temporal integration. Recently, our lab has developed a theoretical framework and general model for describing how these functions are implemented in

the brain. This framework, henceforth referred to as *LeabraTI*,¹ brings together a large number of independent findings from the systems neuroscience literature to describe how multiple interacting mechanisms accomplish prediction and temporal integration in cortex. The fundamental proposal of the LeabraTI theory is that time in the brain is discretized (at least partially) into reference frames that can be associated via general learning mechanisms (e.g., Hebbian and error-driven learning) so that representation of information during one frame can be used to predict what happens during the next. This proposal requires at least two mechanisms: (1.) A mechanism that establishes reference frames over which information is integrated and, (2.) A mechanism that generates the actual prediction itself and validates it against what actually happens.

As will be discussed in detail in Chapter ??, the LeabraTI theory's proposed mechanisms are suggested to be dissociable, generating signatures at distinct spectral frequencies which can be measured physiologically. The potential dissociability of the spatial and temporal components of prediction establishes a number of immediately testable predictions that will form the experimental component of the proposed work.

References

- Busch, N. A., Dubois, J., & VanRullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. *The Journal of Neuroscience*, 29(24), 7869–7876.
- Foldiak, P. (1991). Learning invariance from transformation sequences. *Neural Computation*, 3(2), 194–200.
- Masquelier, T., & Thorpe, S. J. (2007). Unsupervised learning of visual features through spike timing dependent plasticity. *PLoS Computational Biology*, 3(2), 247–257.
- Mathewson, K. E., Lleras, A., Beck, D. M., Fabiani, M., Ro, T., & Gratton, G. (2011). Pulsed out of

¹Leabra refers to a general model of learning in the neocortex (O'Reilly & Munakata, 2000; O'Reilly, Munakata, Frank, Hazy, & Contributors, 2012); TI to Temporal Integration.

awareness: EEG alpha oscillations represent a pulsed-inhibition of ongoing cortical processing. *Frontiers in Psychology*, 2.

Mutch, J., & Lowe, D. (2008). Object class recognition and localization using sparse features with limited receptive fields. *International Journal of Computer Vision*, 80(1), 45–57.

O'Reilly, R. C., & Munakata, Y. (2000). *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain*. Cambridge, MA: The MIT Press.

O'Reilly, R. C., Munakata, Y., Frank, M. J., Hazy, T. E., & Contributors (2012). *Computational Cognitive Neuroscience*. Wiki Book, 1st Edition, URL: <http://ccnbook.colorado.edu>.

Serre, T., Oliva, A., & Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. *Proceedings of the National Academy of Sciences of the United States of America*, 104(15), 6424–6429.

VanRullen, R., Busch, N. A., Drewes, J., & Dubois, J. (2011). Ongoing EEG phase as a trial-by-trial predictor of perceptual and attentional variability. *Frontiers in Psychology*, 2.