Consciousness appears to be a bi-stable phenomenon: you either perceive something, or you don’t. This suggests a simple paradigm for studying the neural correlates of conscious activity: correlate neural activity to perception. This approach, coined the “search for neural correlates of consciousness (NCC)” (Crick and Koch, 1992) has made considerable progress in identifying neural activity that is uncorrelated with consciousness. The problem is that many unexpected processes turn out to be correlated with consciousness, such as attention, behavioral outputs, etc. I propose here to more carefully identify the processes that precede and follow conscious neural activity. For example, attention acts as a gatekeeper of consciousness: when we attend to something we become consciously aware of it. But this could be because attention affects downstream conscious neural processes, or it could be because attention acts directly upon conscious neural processes. To clarify this question I am building a model of feature-based attention as a part of my research program with Prof. Justin Gardner at Stanford University. Our goal is to build a computational model of feature-based attention that can account for the changes that attention causes for conscious visual perception.

**Inattentional Blindness**

In the early 1990s Arien Mack and Irvin Rock discovered that attention to one feature in an image was sufficient to abolish conscious awareness of other features. They coined this phenomena “inattentional blindness” noting in particular that it even occurred when participants were specifically fixating (i.e. their eyes were directly centered on) the feature of which they had no perception. Perhaps surprisingly some features survived inattentional blindness, in particular natural images and faces (Li et al., 2002; Reddy et al., 2004). This led researchers like Christof Koch to announce that attention and conscious awareness were two separable processes in the brain (Koch and Tsuchiya, 2007). More recent research has re-opened the debate by suggesting that even scene perception is subject to inattentional blindness when attention is engaged by a sufficiently difficult task (Cohen et al., 2007). I believe that this mixture of results is in part due to the lack of precision around the process of attention. Attention is a specific neural process. Prof. Gardner has shown that spatial attention to contrast, for example, acts as a selection bias for signals in early visual cortex (Pestilli et al., 2011). If we knew precisely how attention acts on neural processes in visual cortex we would have a better grasp on understanding why attention manipulates our conscious visual perceptions.

**Current and Proposed Studies**

Over the past year I have collected data to begin building a model of the effect of feature-based attention on neural processing in the visual cortex. Using an inattentional blindness paradigm I combined two well understood image features: image contrast and motion coherence. We know that early visual cortical areas such as V1, when measured with BOLD fMRI, show sensitivity to contrast intensity, but not to motion coherence. The reverse is true in the later cortical areas V3a and hMT, which show response sensitivity to motion coherence but not to contrast intensity. Based on this knowledge my prediction was hierarchical: I expected that due to the feed-forward connections in visual cortex any effect of attention would be fed to downstream cortical regions. Specifically, I expected that attention to contrast, which is known to affect V1, would cause corruption or suppression of signals in the downstream areas V3a and hMT. But we also expected that attention to motion, which we expected to affect V3a/hMT would not result in any change in the signals in V1. If discrimination of contrast depended only on V1 and motion only on V3a/hMT, then the behavioral results are clear: attention to motion should affect the perception of contrast but not vice versa. This is precisely the effect that we observed. Our BOLD fMRI results showed that the responses in area hMT and V3a were modulated by attention condition, whereas responses in V1 were not. In parallel we found that discrimination of motion in our task was affected by attention, whereas discrimination of contrast was not.

The model is as yet incomplete—although we know from our data that the perception of contrast and motion are asymmetrically related to attention and that BOLD fMRI responses reflect perception, we still do not know precisely the causal process of attention. My hypothesis is that attention to contrast *corrupts* the representation of motion in the downstream regions. I propose to test this hypothesis in two ways: **1** by building a computational model of our current dataset, looking at what effect of attention best explains our data, and **2** in a new experiment using transcranial magnetic stimulation. TMS is a technique that can be used to either abolish neural activity in a region, akin to a temporary lesion, or to boost or suppress activations that are near perceptual threshold. We can take advantage of both of these techniques to test our hypothesis in the following ways: **1** as a test of causality, using temporary TMS lesions to test the involvement of each area in contrast and motion perception, and **2** by using low-threshold pulses to mimic corrupting attentional signals. If corruption is indeed the cause of our effect than a low-threshold pulse to V1 should introduce similar corruption to downstream regions, mimicking the behavioral effects I have already observed.

**Merit & Impact**

This research has the potential to help clarify the debate surrounding attention and ultimately improve our understanding of the neural processes of consciousness. As I build my model of feature-based attention I continue to think about how this model fits into the larger space of consciousness. Attention is clearly an integral part of our own conscious experiences, but it remains entirely unclear whether the neural process of attention interacts directly with the neural processes of consciousness. Answering that question requires better theoretical models of both attention and consciousness, which will eventually help us understand how conscious neural processes interact with other brain processes. Attention is a good starting point because of the vast psychology literature surrounding inattentional blindness and other related phenomena. Connecting these tasks to human brain anatomy and to computational models is daunting, but the necessary resources and collaborations exist to make this possible, both at Stanford and within my extended network.

**References 1** Crick, F., & Koch, C. (1992). The problem of consciousness. *Scientific American*, *267*(3), 153-159. **2** Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences*, *99*(14), 9596-9601. **3** Reddy, L., Wilken, P., & Koch, C. (2004). Face-gender discrimination is possible in the near-absence of attention. *Journal of Vision*, *4*(2), 4. **4** Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends in cognitive sciences*, *11*(1), 16-22. **5** Cohen, M. A., Alvarez, G. A., & Nakayama, K. (2011). Natural-scene perception requires attention. *Psychological science*. **6** Pestilli, F., Carrasco, M., Heeger, D. J., & Gardner, J. L. (2011). Attentional enhancement via selection and pooling of early sensory responses in human visual cortex. *Neuron*, *72*(5), 832-846.