Psych 202: Paper #1

Birman

**Psychophysical Functions as Incomplete Models of Subjective Visibility**

A fundamental issue for our understanding of consciousness is whether awareness is a graded continuum or all-or-none. These different hypotheses would require vastly different implementations in cortex: graded awareness implies a neural mechanism that scales, while an all-or-none threshold suggests a separate network might “turn on” during awareness. A number of experiments have found contradictory evidence for either a graded or threshold account of visual awareness. In a series of studies Overgaard and colleagues found evidence for the graded awareness of stimuli, based on gradual increases in performance with increasing stimulus strength and gradual increases in subjective reports of awareness (for an overview see Sandberg, Bibby, Timmermans, Cleeremans, & Overgaard, 2011). Neuroimaging evidence appears to corroborate a graded account for some stimuli such as shapes (Christensen, Ramsøy, Lund, Madsen, & Rowe, 2006). In contrast a number of phenomena show sudden transitions into awareness, for example in the perception of Mooney images, finding solutions to arithmetic problems, and the detection of gist in natural scenes. A plausible explanation for these differences is that these stimuli are represented in different ways, leading to different perceptual awareness. The stimuli in experiments that report a graded increase in awareness are consistently basic visual features: e.g. contrast, color, and orientation. In contrast Mooney images often use faces, which, along with numbers and natural scenes are processed and represented by higher stages in the visual processing hierarchy (Rees, Kreiman, & Koch, 2002). Windey et al. (Windey, Gevers, & Cleeremans, 2013) propose that this distinction between earlier and later stages of the visual processing hierarchy is what differentiates continuous transitions into awareness from distinct thresholds. This hypothesis additionally suggests different underlying neural correlates of consciousness (NCC) at different stages of the visual hierarchy, a finding that should cause a shift in the focus of the search for NCCs. Although the hypothesis and modeling approach of Windey et al. are an exciting addition to the available toolkit for studying visual awareness their methodology does not warrant their conclusions. In this brief commentary I outline a methodological issue with their approach and steps that could be taken to improve our understanding of this interesting and exciting question.

Windey et al. set out to induce different processing in the visual hierarchy from identical visual stimuli. To produce this effect they presented subjects with consistent stimuli but varied the two-alternative task they were asked to perform to be either color discrimination or value judgment. Subjects in their experiments were presented with red or blue colored numbers from 1-9 for 10-80 ms, forward and backward masked by an image consisting of many squares of red and blue color. During two response periods participants reported the color (red or blue) or value (> or < 5) and then their subjective awareness on a 1-4 scale (the Perceptual Awareness Scale, Overgaard et al. 2006). The performance data was parameterized with a model of the Weibull function, where two parameters indicate the upper and lower bounds of performance (floor and ceiling) and two parameters indicate the inflexion point (stimulus strength threshold for 75% performance) and the slope. They found that the slope parameter *d* differed between tasks such that the slope was steeper in the “high-level” task involving numerical judgments; *d*\_low = 2.401 (SE = 0.174), *d*\_high = 1.969 258 (SE = 0.122), t(19) = 2.033, p = 0.028 (Windey et al., 2013). They did not find significant differences in performance between tasks at any stimulus strength. The authors suggest this indicates that level of processing was successfully manipulated without any change in task difficulty. They interpret their results to mean that taking into account level of processing is sufficient to integrate the conflicting psychophysical results reviewed earlier.

Although the results reported by Windey et al. are compelling and suggest a significant breakthrough in a major issue facing our understanding of perceptual awareness, they are not as strong as the authors imply. Their design employs a standard psychology approach in which a number of college-age students were asked to perform a simple task for ~1 hour. Each participant therefore saw 32 trials per task corresponding to 4 examples per stimulus strength. The estimates of *d* (the slope parameter) were computed per-subject, averaged, and compared through a *t-test* between tasks. This design and analysis are inappropriate for the question that was posed by the authors. One major issue is related to the estimation of the *d* (slope) parameter, which has a history of being problematic in psychophysics (see for example Leek, 2001).

To understand the psychophysics approach employed by Windey et al. it helps to be grounded in the history of modeling discrimination tasks. Discrimination tasks like the one used by Windey et al. are a variant of signal detection tasks (Swets, Tanner Jr, & Birdsall, 1961). Instead of asking participants to report a “yes-no” response for detection they are asked to report about two alternatives of a stimulus feature. This forces subjects to guess when no information is available, effectively pushing their criterion to 0. In signal detection tasks on the other hand shifts in criterion imposed by . The results of two-alternative experiments are usually plotted as functions of measured performance against stimulus threshold, to which a parameterized function is fit (as above). According to Swets et al. we can imagine the two-alternative discrimination task to be a choice between Gaussian distributions along a stimulus feature dimension or two dimensions each with a signal and noise. An observation then is categorized according to some criterion as coming from one of the two signal distributions where the criterion maximizes the rate of success on the task. One peculiar issue that continues to plague signal detection and psychophysics in general is the question of how to estimate the underlying distributions and their parameters. One approach, and that used by Windey et al, has been to parameterize the performance functions and compare parameters across tasks. This approach is problematic: estimates of parameters are known to be skewed and comparisons are therefore best done via bootstrapping (Maloney, 1990; Wichmann & Hill, 2001). In addition, estimating a slope requires significantly more trials than estimating a threshold (a specific stimulus->performance mapping, e.g. the stimulus strength corresponding to 70% performance) (King-Smith, Ewen P. & Rose, David, 1997). King-Smith & Rose report that with 50 trials the variance of their slope parameter remains large, more than an order of magnitude larger than the threshold, which stabilizes within 50 trials. Kontsevich & Tyler report similar results, that within 30 trials the threshold is precise to 2 dB (23%) but it takes 300 trials to achieve similar precision for the slope (Kontsevich & Tyler, 1999). Accordingly they outline an adaptive procedure to efficiently collect slope data. Note that Windey et al. recorded 32 trials per task for each participant, corresponding to four trials at each of 8 stimulus strengths. Their choice was undoubtedly motivated by cost and convenience, but according to the brief review of slope estimation outlined above it remains entirely incompatible with their analysis.

One could argue that by averaging across subjects they avoid the need to accurately estimate the slope within individuals, which leads us to the second major issue with this approach. Slope estimates are a parameterization of an underlying physical phenomenon. The assumption is that in a given brain processing that leads to consciousness of low-level features is graded, while processing for high-level features is some threshold function of lower-level activation. But this difference may be inconsistent across brains showing large variability. Windey et al. do not report any information about this, but based on the results of Kontsevich et al. we can estimate that even within a subject slope estimates will show huge variability with only 32 trials recorded. The difference in *d* as a within subject variable was also not reported, although it would be the correct measurement to make since we expect variability in *d* to occur within brains, where each subject may have additional *d* variance across the population.

Understanding the nature of conscious experience is a difficult and contentious goal and small steps like those outlined by Windey et al. are of the utmost importance. By parameterizing the psychometric function of performance for tasks known to rely on activation of different parts of the visual hierarchy, but using similar base stimuli, they make an important contribution to our understanding of how visual perception occurs in the human brain. Nevertheless, their contribution is reduced by their use of mismatched computational methods relative to their proposed goals. Signal detection methodologies are clearly well matched to detecting changes in consciousness and the parameters that influence conscious perception. But they also require a careful understanding of the underlying assumptions that these cognitive models were designed for, which may be incompatible with certain experimental designs. I will sketch here the changes necessary to match the paradigm of Windey et al. with the modeling techniques used in signal detection theory.

**References**

Christensen, M. S., Ramsøy, T. Z., Lund, T. E., Madsen, K. H., & Rowe, J. B. (2006). An fMRI study of the neural correlates of graded visual perception. *NeuroImage*, *31*(4), 1711–1725. http://doi.org/10.1016/j.neuroimage.2006.02.023

King-Smith, Ewen P., & Rose, David. (1997). Principles of an Adaptive Method for Measuring the Slope of the Psychometric Function. *Vision Research*, *37*(12), 1595–1604. http://doi.org/10.1016/S0042-6989(96)00310-0

Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research*, *39*(16), 2729–2737.

Leek, M. R. (2001). Adaptive procedures in psychophysical research. *Perception & Psychophysics*, *63*(8), 1279–1292.

Maloney, L. T. (1990). Confidence intervals for the parameters of psychometric functions. *Perception & Psychophysics*, *47*(2), 127–134.

Rees, G., Kreiman, G., & Koch, C. (2002). NEURAL CORRELATES OF CONSCIOUSNESS IN HUMANS. *Nature Reviews Neuroscience*, *3*(4), 261–270. http://doi.org/10.1038/nrn783

Sandberg, K., Bibby, B. M., Timmermans, B., Cleeremans, A., & Overgaard, M. (2011). Measuring consciousness: Task accuracy and awareness as sigmoid functions of stimulus duration. *Consciousness and Cognition*, *20*(4), 1659–1675. http://doi.org/10.1016/j.concog.2011.09.002

Swets, J. A., Tanner Jr, W. P., & Birdsall, T. G. (1961). Decision processes in perception. *Psychological Review*, *68*(5), 301.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception & Psychophysics*, *63*(8), 1314–1329.

Windey, B., Gevers, W., & Cleeremans, A. (2013). Subjective visibility depends on level of processing. *Cognition*, *129*(2), 404–409. http://doi.org/10.1016/j.cognition.2013.07.012