Psych 202: Paper #1

Birman

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

A fundamental issue for our understanding of consciousness is whether conscious percepts are graded or all-or-none phenomena. The implication of this difference suggests dramatically different implementations in cortex: a gradation of awareness implies a neural mechanism that scales accordingly, while an all-or-none threshold suggests a separate network that “turns 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. In addition, neuroimaging evidence appears to corroborate this graded account. In contrast a number of phenomena show sudden transitions into awareness, for example in Mooney images, solutions to arithmetic, and natural scene detection. One possible explanation for these differences lies in their neural substrates. The stimuli in experiments that report a graded increase in awareness are basic visual features: e.g. contrast, color, and orientation. In contrast Mooney images often use faces which, along with numbers and natural scenes, are known to be processed and represented by higher stages in the visual processing hierarchy (CITE). 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 is an exciting addition to our understanding of consciousness the methods presented by Windey et al. do not warrant their conclusions. In this brief commentary I outline methodological changes that could help answer this interesting and exciting question.

To induce stimulus processing at different stages of the visual hierarchy Windey et al. employ a classic two-alternative discrimination task. Subjects in their experiments were presented with constant stimuli: red or blue letters from 1-9 for 10-80 ms, forward and backward masked by an image consisting of many squares of red and blue color. In their experiment participants either attended to the color feature of the stimulus or to the numerical value. 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 analyzed with a 4-parameter model of a 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 in the slope was steeper in the “high-level” task involve 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 at any stimulus strength between their tasks which they suggest indicates that level of processing was successfully manipulated without any change in task difficulty. They interpret their results as showing 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 of each task corresponding to 4 examples at each stimulus strength. The estimates of *d* (the slope parameter) were computed per-subject, averaged, and compared through a *t-test* between participants. This design and analysis are inappropriate for the question that was posed by the authors. One issue is related to the estimation of the *d* (slope) parameter, which has a history of being a problematic issue in psychophysics, see for example (Leek, 2001). A second issue relates to their problematic assumption that difficulty has been maintained across tasks, which is in part related to their choice of doing an across-subject analysis instead of a within-subject analysis. Below I discuss these two major issues and how they can be avoided as well as an additional modeling analysis that could provide significant insight into the processes at play.

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 specific stimulus feature. This forces subjects to guess when no information was available, effectively pushing their criterion to 0. This has the major advantage of being a criterion-free measurement, which would otherwise be a confound of awareness. 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 which 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. This is a within-brain difference, but these assumptions make no

that processing level implies physical constraints that lead to either graded or threshold awareness we should be careful to specify the assumptions that we need to make about

**References**

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.

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