**Metacognition: Using confidence ratings for Type 2 and Type 1 ROC curves.**

Stanley Klein, UC Berkeley

There is some confusion about how to make use of confidence ratings in a wide variety of experiments. The typical experiment involves a binary judgment ("I saw it or remembered it") followed by a multilevel confidence rating. It is a confusing topic because the rating could be based either on confidence of one's response (standard rating Type 1 ROC) or on confidence sorted on whether that response was correct (Type 2 ROC). Both of these situations can be called metacognition since in both cases the confidence rating is a "meta" judgment about one's own decision process (thinking about thinking). This "meta" field of study has recently become quite popular as can be seen in the recently published book edited by Stephen Fleming and Christopher Frith: "The Cognitive Neuroscience of Metacognition". It is a compilation of 17 survey articles on the topic. The publisher, Springer, has made the electronic version of the book freely available to the public (see below).

My overview of this topic will have three parts:
1) A brief overview of Type 2 ROC based largely on the recent work by Maniscalco and Lau (2012, Fleming book chapter 3), and by Fleming et al. (Fleming book chapters 1 and 11) show that it is fully relevant to vision as well. Another excellent overview is by Galvin, et al. (2003). The Fleming/Frith book makes it clear that one of the most important fields of study that is relevant to meta is the field of memory research. I will argue that Type 2 ROC metacognition has minimal relevance to visual psychophysics. It is a confusing topic largely because the underlying distributions on which the ROC curve is based are far from being Gaussian.

2) There have been interesting controversies in memory research related to whether source memory has a Gaussian distribution or whether it belongs to high threshold processing and ROC curves. Fifteen years ago my graduate student, Scott Slotnick and I joined Art Shimamura and Chad Dodson in a fascinating memory study (Slotnick, et al., 2000) "An Analysis of Signal Detection and Threshold Models of Source Memory" (2000) The task involved memorizing a number of items. Then in the recall part the subject gave a triple judgment: a) "old" vs "new", b) A confidence rating of 1-6 on "old" or "new" and c) a source rating of 1-6 ranging from female to male regarding the gender of who stated the word if it was 'old'. The interesting controversies are related the underlying statistics about the judgments. Three theories were compared, signal detection theory, multinomial modeling and double high threshold theory. The controversy about the underlying detection models are given in the articles by Slotnick et al. (2000) and Yonalinas (1999).

3) An interesting question regarding Type 1 ROC analysis is how does the d' depend on confidence. A recent study on that topic was by Wolfe and Whitney (2014) on discriminating the emotion of a crowded face in peripheral vision. A binary choice was made followed by a confidence judgment. The interesting manipulation was in comparing two types of attention: endogenous vs presaccadic. The fascinating result for their best subject was that for the presaccadic case, the d' was significantly lower at low confidence (note that this point isn't visible in their plots). It is important to realize that in standard binary judgments the criterion is placed at the low confidence point. Thus standard, two response psychophysics may be underestimating the discriminability of stimuli. Note that the meta methodology with just two stimuli plus a binary judgment followed by a rating, is quite different from a rating scale methodology where there are as many stimuli as there are responses and there could well be subtle sensory cues that influence the ratings.

**References.**

Galvin SJ, Podd JV, DRGA V, Whitmore J (2003). Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions. *Psychonomic Bulletin & Review*. 10(4) 843-876.

Fleming S, Frith (2014) Metacognition. http://download.springer.com/static/pdf/630/bok%253A978-3-642-45190-4.pdf?auth66=1425119491_a4dd1e109ae88e7d755a6465bf2d0862&ext=.pdf

Maniscalco B, Lau, H (2014). Signal detection theory analysis of type 1 and type 2 data: meta-d', response-specific meta-d', and the unequal variance SDT mode. In S. M. Fleming & C. D. Frith (Eds.), The Cognitive Neuroscience of Metacognition (pp.25-66). *Springer*.

Slotnick SD, Klein SA, Dodson CS, Shimamura AP (2000). An analysis of signal detection and threshold models of source memory. *Exp. Psychol. Learn. Mem. Cogn.* 26(6):1499-1517.

Wolfe B, Whitney D (2014). Facilitating recognition of crowded faces with presaccadic attention. *Frontiers in Human Neuroscience*. 8 (103): 1-9.

Yonelinas, AP (1999). The contribution of recollection and familiarity to recognition and source-memory judgments: A formal dual-process model and an analysis of receiver operating characteristics. Journal of Experimental Psychology: Learning, Memory, and Cognition, 25, 1415- 1434.