Final accepted version of:
Dube, B. and Golomb, J.D. (2020). Revisiting mixture models of memory. *Nature Human Behavior.* 4: 1098-99. https://doi.org/10.1038/s41562-020-00947-z.

**Title:**
Revisiting mixture models of memory

**Authors**:
Blaire Dube and Julie D. Golomb
Department of Psychology, The Ohio State University
(Correspondence to golomb.9@osu.edu)

**Subject strapline:**
Visual working memory, mixture model

**Standfirst:**
Probabilistic mixture models have contributed significantly to advancements in visual working memory research in recent decades. In a recent paper, Schurgin and colleagues revisit the basic assumptions of mixture models, and suggest that we cannot understand memory without first considering perception.

**Main text:**

Visual working memory—our short-term storage system for visual information—is a core cognitive function foundational to human behavior. Its broad relevance has made visual working memory a hotbed for research and debate, with researchers proposing competing theories describing its capacity limitations, the architecture that defines it, and the intricacies of its relationship with attention, perception, intelligence, and cognitive control[1,2]. In their recent Nature Human Behavior paper, Schurgin, Wixted, and Brady[3] force us to revisit some core assumptions of visual working memory and perception, and the tools we use to study them.

Over the past two decades, a popular tool has emerged for measuring various aspects of visual working memory performance: probabilistic mixture models[4]. A typical mixture modeling experiment is as follows: A participant memorizes the colors of four squares, and after a few-second delay, is cued to report the color they remember seeing in the cued location, by clicking on a color wheel. This continuous report technique produces a distribution of memory responses that allows for more nuanced information about memory than can be obtained from discrete correct/incorrect responses. The essence of mixture modeling is that one takes this distribution of responses and fits it with a probabilistic model representing a mixture of different types of responses. The most common mixture models used to study visual working memory assume two distinct memory states: remembered items (which should produce responses clustered around the correct value, with a flexible parameter for precision) and forgotten items (which should produce random guessing across all values uniformly). Thus, the mixture modeling technique aims to differentiate guesses from correct responses (and sometimes additional types of errors), and measures memory precision and guess rate as separate—and meaningful—parameters. An entire sub-field has been built around how these parameters vary with different task demands, types of memory, and individual and clinical differences.

Schurgin et al. claim that extant models of visual working memory relying on this technique are flawed in that they do not consider the psychophysical similarity of memory items. Using color space as an example, they demonstrate that the perceptual similarity between two colors is not simply a linear function of their distance along the color wheel. For example, two colors 5deg apart in color space are perceived as much more similar than two colors 35deg apart in color space. But, two colors 120deg apart in color space are *not* perceived as more similar than two colors 180deg apart in color space. The authors' critical claim is that when psychophysical similarity is considered, working memory performance can be boiled down to a noisy familiarity signal, with the target color—and those that are close to it in feature space—benefitting from a bolstered familiarity signal that drops off and evens out with increasing distance. This leads to a bold claim: that working memory performance can be explained by a relatively simple signal detection framework with only a single free parameter. In other words, precision and guessing are not distinct kinds of errors, but arise from a single process where "memory strength" (d-prime) is the only memory-based parameter that varies. The authors go on to show that this new Target Confusability Competition (TCC) model accounts for a wide variety of results, including data newly collected as well as existing data from the literature, manipulating set size, encoding times, and delays. They also demonstrate that this psychophysical scaling function can be measured from one set of participants and then applied to data from other participants, that the TCC model can be generalized to other stimulus spaces, including faces, and can generalize across both working memory and long-term memory tasks, with only the critical d-prime parameter varying to account for the results.

The authors' proposal is provocative, particularly in their assertion that there is no objective "guessing" in working memory, a claim which is intuitively hard to swallow. Yet their paper is also undeniably sophisticated and thorough: The sheer volume of experiments that they include is impressive, and the evidence that they present, bolstered with intuitive explanations of how a single d-prime parameter could account for our subjective sense of things like guessing and confidence, are convincing and difficult to argue with—even, perhaps, for the biggest skeptics among us.

If the authors are correct, what does this mean for the field? Mixture models, and the assumptions that underlie their use, have become a major part of the visual working memory (and broader) literature. If guessing and memory precision are not distinct psychological states, the greatest impact will likely be on studies that aim to make meaningful claims based on independent fluctuations in the standard deviation and guess rate parameters of a mixture model. But it is important to keep in mind that many of the insights that we have gained from observing memory performance under different experimental conditions remain valid. Moreover, in many applications, mixture models have evolved from the basic two parameter precision-plus-guessing models to incorporate other types of errors, such as swaps and shift errors[5,6]. Though the TCC model in its current form is attractive in its simplicity, it is essentially a new base model of memory: in future work it could and arguably should be adapted from a single-parameter model to incorporate some of these additional parameters. It will be interesting to see how both the model and related theoretical questions evolve as a result.

Whether or not the bold claims of Schurgin et al. turn out to be correct, science benefits from researchers who are willing to question foundational assumptions. With their relatively simple TCC model, Schurgin et al. force us to revisit the very basics of a field, an approach that is both admirable and worthy of our attention—especially in light of the increasing tendency for more and more complex models. There will likely be opposition to some of their claims, but their paper should not be ignored, and whatever debate ensues will undoubtedly advance our understanding of visual working memory and the important relationship between memory and perception.

**References**

1. Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, 17, 391-400.

2. Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience,* 17, 347-356.

3. Schurgin, M. W., Wixted, J. T., & Brady, T. F. (2020). Psychophysical scaling reveals a unified theory of visual memory strength. *Nature Human Behavior*.

4. Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, 453, 233-235.

5. Bays, P. M., Catalao, R. F., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, 9, 7-11.

6. Golomb, J. D., L'Heureux, Z. E., & Kanwisher, N. (2014). Feature-binding errors after eye movements and shifts of attention. *Psychological Science*, 25, 1067-1078.

**Competing Interests:**
The authors declare no competing interests.