Precision-weighted evidence integration predicts time-varying influence of memory on perceptual decisions

Ari Khoudary (ari.khoudary@uci.edu)

Megan A. K. Peters* (megan.peters@uci.edu)

Aaron M. Bornstein* (aaron.bornstein@uci.edu)
Department of Cognitive Sciences, University of California, Irvine

Abstract

How do we use past experiences to make sense of the present? We developed an experimental task and computational model to investigate whether human observers Bayes-optimally integrate evidence from memory and vision during cued perceptual decisions. Drawing on theory developed in the multisensory integration literature, we model the decision process as precision-weighted integration of samples from memory and vision, with the rate of accumulation defined by the relative precision of each evidence stream. The model captures two qualitatively distinct empirical findings from different task structures, suggesting this framework has the potential to identify a process fundamental to decisions relying on multiple evidence sources. The experimental task will measure how human observers actually perform this integration, which will allow us to better characterize the conditions under which humans deviate from optimal choice behavior.

Keywords: evidence accumulation; perceptual decisions; memory-guided behavior; uncertainty; simulation

Introduction

The present work synthesizes findings from two previous studies investigating the dynamic integration of prior expectations into perceptual inference. The first study examined the integration of priors over the identity and coherence of visual information that were signaled trial-by-trial by fractal cues (Bornstein et al., 2018). Participants in this study both used the cue to retrieve mnemonic evidence of the anticipated upcoming visual percept, and simultaneously modulated their reliance on the cue according to the signaled coherence of the upcoming visual evidence stream. Specifically, fMRI evidence of memory retrieval before the onset of visual information was lower when the cue signaled that upcoming visual information would be of higher coherence. Separately, another study examined how humans and non-human primates integrated prior expectations about the direction of random moving dots of unknown coherence (Hanks, Mazurek, Kiani, Hopp, & Shadlen, 2011). In this study, the priors about direction were established in blocks across experience, rather than signaled at each trial. The researchers observed that, as time within each trial passed, participants' behavior and neural activity evidenced a stronger influence of these prior expectations. They proposed that participants estimated coherence dynamically, using a 'Time-Dependent Accuracy' function, and mixed in the prior in proportion to elapsed time as a proxy for the coherence of the underlying visual evidence.

Here, we develop a computational framework that predicts the moment-by-moment timecourse of the decision variable during cue-guided perceptual decisions as it would result from Bayes-optimal integration of two overlapping noisy evidence streams, following relevant work on multisensory integration (Angelaki, Gu, & DeAngelis, 2009). The mathematical framework we develop explains these two results using a single normative process (Wang, Feng, & Bornstein, 2022). Specifically, we propose that choices are the result of a precision-weighted integration of samples from each evidence stream, where precision is first anticipated by either cued information or cross-trial average coherence and then, as dynamic evidence becomes available, estimated in real time as the relative inverse entropy of each unfolding stream. This framework provides a principled explanation of time-varying drift rates across a wide variety of experimental conditions, and establishes a tool for examining multi-modal integration of dynamic evidence streams more broadly.

Task

Figure 1 depicts the task designed to be completed by human observers. Participants learn, through experience, probabilistic cue-target pairings and use these memories to complete a perceptual inference task. There are two categories of targets (faces and houses), each consisting of two perceptually-similar grayscale images. Categories are assigned to either strong (80%) or weak cue (65%) strength. Cues consist of fractal images, and each target within a category is associated with a unique cue. Additionally, each category has a chance (50%) cue; these are learned explicitly during the instruction phase of the task (Fig. 1A). We refer to these cue strength values as memory evidence reliability. During *Cue Learning* (Fig. 1B), participants are presented with a cue and, after a brief interstimulus interval (ISI), one of the target images from the appropriate category. Their task is to press the button that indicates the target's identity. During *Cued Inference* (Fig. 1C), memory sampling is initiated by presenting participants with a fractal cue. After a variable ISI, visual sampling is initiated by the onset of a "flickering stream" of same-category images. Participants' task is to determine which of the images was presented more frequently in the stream (i.e., is the "target") and press the button that signals its identity. *Coherence* is defined as the proportion of frames that include the target image, and this value is set pseudorandomly on each trial to be 50%, 65%, or 80% (corresponding to chance, weak, and strong visual evidence).

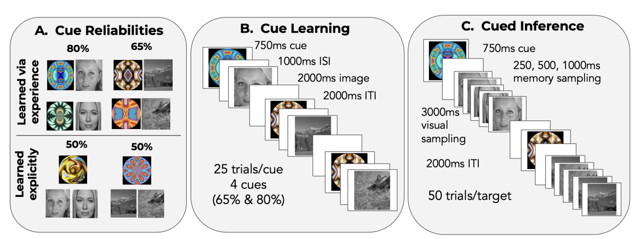

Figure 1: Task design.

Precision-weighted multi-stage evidence accumulation model

The model is an extension of a time-varying drift diffusion model developed previously (Srivastava, Feng, Cohen, Leonard, & Shenhav, 2017). At each time point, evidence samples are drawn from a binomial distribution with probability of success (*target* image) defined by cue strength (for memory) or coherence (for vision), with additive zero-centered Gaussian noise (Eq. 1). This sample is used to update the probability of observing a target (Eq. 2), which is then used to compute time-varying precision as the inverse entropy of the evidence stream (Eq. 3). Next, we define each evidence stream's drift rate as the relative precision between the two streams (Eq. 4), and use this to compute a weighted sum of the two samples to serve as 'evidence' at that timestep (Eq. 5). This process continues until the summed evidence passes a fixed threshold, at which point a decision is made. Formally,

$$o_{memory} \sim Binomial(cueStrength, 1) + Normal(0, 1) \tag{1}$$

$$p(target|evidence_{memory})_t = \frac{\sum_{t=1}^{nSamples} o_{memory} > 0}{nSamples} \tag{2}$$

$$memoryPrecision_t = H(p(target|evidence_{memory})_t)^{-1} \tag{3}$$

$$memoryDriftRate_t = \frac{memoryPrecision_t}{memoryPrecision_t + visualPrecision_t} \tag{4}$$

$$\begin{aligned} DV_t = DV_{t-1} + o_{memory} * memoryDriftRate + \\ o_{vision} * visualDriftRate \end{aligned} \tag{5}$$

where H is the Shannon entropy function and DV is the decision variable. For brevity, we omitted the formulas used to compute parallel values for vision, in which case all instances of *memory* above can be swapped with *visual*, and for Eq. 1 *coherence* is used in place of *cueStrength*.

Simulation details

We examined the timecourse of the decision variable under conditions of varying prior information about target identity and coherence. Each run of the simulation consisted of 25 trials for 100 participants. On each trial 100 visual samples and 125 memory samples (25 during the pre-flicker "retrieval" period, 100 during visual sampling). Values for trial number and number of samples in each evidence stream were chosen to match the timing of future experiments involving human subjects. The MATLAB code is available at https://github.com/ari-khoudary/time-varying-precision-weighted-msddm.

Model assumptions

To compute precision-weighted drift rate during the retrieval period when no visual evidence was available, we defined visual precision as $H(0.8)^{-1}$. This was motivated by Bornstein et al.'s (2018) finding that memory drift rate is lower when the cue indicates that upcoming visual evidence would be of high coherence, a key explanatory target of this model. For the first sample of visual evidence, precision was defined as $H(0.65)^{-1}$. All other precision computations followed equation 3 and its visual analog. Accumulated memory evidence and precision estimates were carried forward from the retrieval period to memory sampling during the flicker stream.

Results & discussion

As shown in Figure 2 (top), when visual evidence is noisy, choice is driven by memory evidence at a rate proportional to the precision of that memory evidence. Note that the *relative* precision of memory increases as visual evidence is progressively determined to be imprecise, across the timecourse of the task, matching the findings of Hanks & colleagues (2011). Figure 2 (bottom) shows that strong memory evidence improves decision efficiency for weakly coherent visual information under conditions of congruency (left), and determines choice under conditions of incongruence (right). In all cases, anticipating high coherence visual information attenuates memory drift rate during the memory retrieval period (time=0:25) relative to the simultaneous visual and memory sampling period (time=26:100).

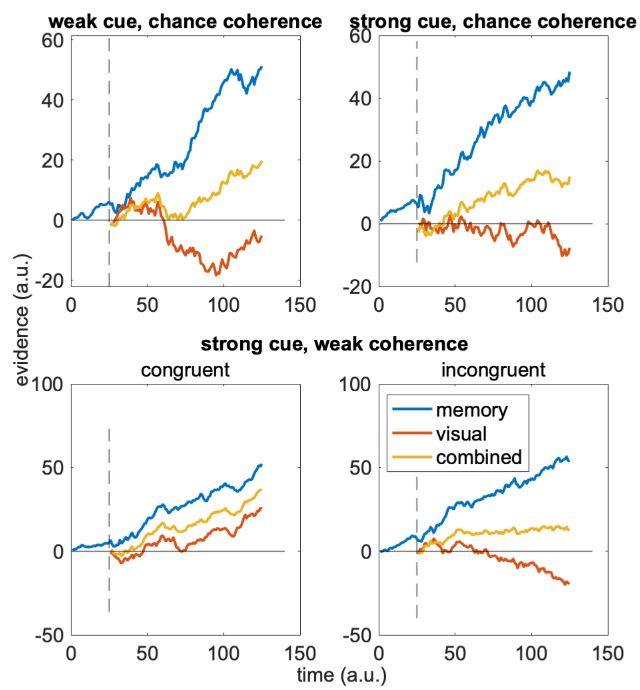

Figure 2: Representative single-trial timecourses. Top: for chance (0.5) coherence visual evidence coupled with weak (0.65; left) and strong (0.8; right) memory evidence. Bottom: for strong (0.8) memory evidence coupled with weak (0.65) visual evidence that is congruent (left) or incongruent (right) with evidence from memory. Vertical dashed line indicates the onset of visual evidence.

Conclusion and future directions

Our model successfully synthesized and reproduced results from previous work investigating how expectations are integrated with ongoing experience to drive perceptual decisions. Future work will expand the experimental design to add trials that measure how human observers estimate the reliabilities of each evidence stream, and investigate how these estimates modulate the Bayes-optimal combination of information from each stream.

References

Angelaki, D. E., Gu, Y., & DeAngelis, G. C. (2009, aug). Multisensory integration: psychophysics, neurophysiology, and computation. *Curr. Opin. Neurobiol.*, *19*(4), 452–458.

Bornstein, A. M., Aly, M., Feng, S. F., Turk-Browne, N. B., Norman, K. A., & Cohen, J. D. (2018). Perceptual decisions result from the continuous accumulation of memory and sensory evidence. *BioRxiv*, 186817.

Hanks, T. D., Mazurek, M. E., Kiani, R., Hopp, E., & Shadlen, M. N. (2011). Elapsed decision time affects the weighting of prior probability in a perceptual decision task. *Journal of Neuroscience*, *31*(17), 6339–6352.

Srivastava, V., Feng, S. F., Cohen, J. D., Leonard, N. E., & Shenhav, A. (2017, apr). A martingale analysis of first passage times of time-dependent wiener diffusion models. *J. Math. Psychol.*, *77*, 94–110.

Wang, S., Feng, S. F., & Bornstein, A. M. (2022). Mixing memory and desire: How memory reactivation supports deliberative decision-making. *Wiley Interdisciplinary Reviews: Cognitive Science*, *13*(2), e1581.