Comparing primacy effects across different temporal scales to distinguish between theories of evidence integration

Ankani Chattoraj[1,⋆], **Xuan Wen**[2], **Ralf M. Haefner**[3,⋆]
1 NVIDIA Corporation, 2788 San tomas Expressway, Santa Clara CA 95051
2 Vanderbilt Brain Institute, Vanderbilt University, 2201 West End Ave, Nashville, TN 37235
3 Brain and Cognitive Sciences, University of Rochester, Rochester, NY 14627, USA
⋆ ankanichattoraj91@gmail.com, rhaefne2@ur.rochester.edu

Abstract

Integrating noisy sensory evidence over time is an essential feature of perception, and of perceptual decision-making. While ideal observer analyses have guided our understanding of its algorithmic and neural basis, behavior often deviates from optimality by being biased towards early evidence. Conflicting theories have been proposed to explain such a primacy effect. Algorithmic level theories include the integration of evidence only until an internal bound is reached, and a confirmation bias due to approximate hierarchical inference. Biophysical theories on Marr's implementation level include neural adaptation and attractor dynamics. Here, we tested these different hypotheses by measuring the strength of the primacy effect in a visual discrimination task across a range of time scales from total stimulus durations of 416ms to 1667ms, and for a range of evidence frame durations from 17ms to 83ms. We find the primacy effect to be invariant with respect to stimulus time, but not physical time. Furthermore, a lower influence of early evidence frames when they are very short argues against neural adaptation. Together, our results favor an algorithmic explanation as proposed by approximate hierarchical inference.

Keywords: evidence integration, perceptual decision-making, primacy, confirmation bias, approximate inference

Introduction: Temporal biases are widely observed in perceptual decision-making tasks, ranging from the overweighting of early evidence ("primacy effect", e.g. (Nienborg & Cumming, 2009; Kiani, Hanks, & Shadlen, 2008; Yates, Park, Katz, Pillow, & Huk, 2017; Lange, Chattoraj, Beck, Yates, & Haefner, 2021)) to the overweighting of recent evidence ("recency effect", e.g. (Drugowitsch, Wyart, Devauchelle, & Koechlin, 2016; Glaze, Filipowicz, Kable, Balasubramanian, & Gold, 2018)) even in tasks when the optimal strategy demands an equal weighting of all evidence frames (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006; Gold & Shadlen, 2007). While compelling explanations exist for the recency effect in terms of expected nonstationarities in the environment (Glaze et al., 2018), and a simple forgetting due to accumulating internal noise, the cause for the observed primacy effects is far less clear. Broadly, the four previously proposed models can be classified into biophysical and algorithmic models. Under biophysical models we include those that explain primacy as the result of neural adaptation, i.e. a decrease in tuning gain over time within a trial (Yates et al., 2017), and those that explain it as the result of attractor dynamics in a decision area with a fixed time constant (Wang, 2002). One algorithmic explanation that has been proposed suggests that later evidence is given less weight *on average across trials* since internal evidence accumulation stops and later evidence is ignored, when a certain bound is reached, even when the stimulus duration is fixed (Kiani et al., 2008). Alternatively, it has been proposed that the primacy effect is the consequence of a perceptual confirmation bias: when performing hierarchical approximate inference the brain's top-down expectations influence lower sensory representations which in turn update task-related beliefs (Lange et al., 2021).

Predictions: Importantly, these hypotheses differ in their predictions about how the primacy effect should scale with stimulus duration. Most simply, models that rely on fixed biophysical processes predict a bias that is a function of physical time (e.g. a decaying exponential), regardless of the duration of the stimulus, while algorithmic models predict a bias that is a function of "computational time", e.g. the number of the stimulus frame presented, not physical time. Similarly, the primacy effect predicted by neural adaptation should be independent of the duration of individual stimulus frames, strictly decreasing over the course of physical time in agreement with the time course of neural gain.

Experimental design & analysis: Here, we present new data, collected during two experiments involving 10 human observers each. In both experiments, observers had to report the perceive orientation of bandpass-filtered Gaussian noise, presented in a sequence of stimulus frames (Fig. 1A). All tasks, and task-related stimuli, were identical except for the duration of each frame, and their number within a trial (Fig. 1B). In the **first experiment**, we compared the effect of three different frame durations *while keeping the total number of frames constant*. In the **second experiment**, we compared the effect of three different numbers of frames *while keeping the total stimulus duration constant*. We collected 60h worth of psychophysical data (10 observers for each of 2 experiment, 3 sessions per observer). We quantified the weight given to each evidence frame using logistic regression using two kinds of regularization: (1) using a cross-validated smoothness constraint (Fig. 2A+C), and (2) using a two-parameter function (pdf of a $\Gamma-$distribution, Fig. 2B+D) that yielded the same qualitative results. Our central comparison consists in the agreement between the three weight profiles when plotted

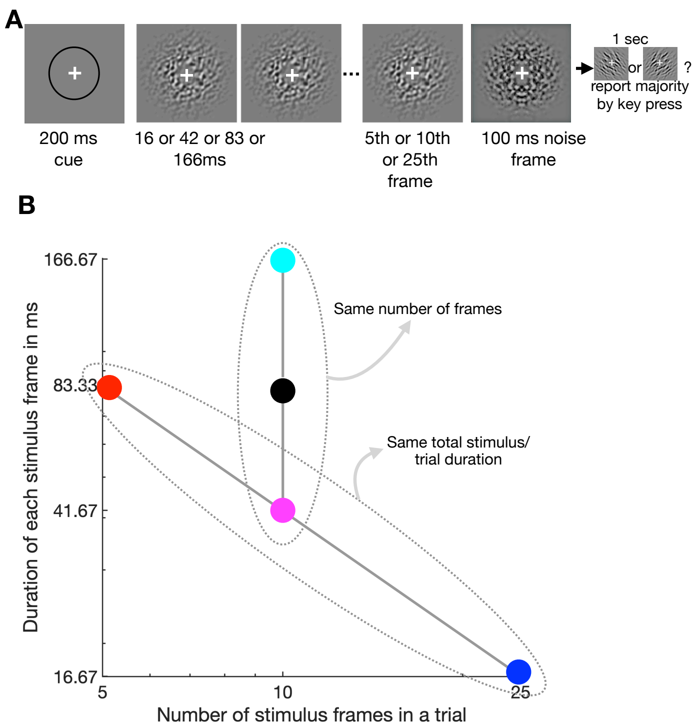

Figure 1: **A:** Evidence integration task **B:** Experimental design to test properties of perceptual confirmation bias

as a function of evidence frames ("algorithmic time", top row in Fig. 2) vs their agreement when plotted as a function of physical time (bottom row in Fig. 2).

Empirical results: First, we find that the perceptual primacy effect is remarkably robust across different time scales from stimulus durations of 416ms to 1667ms, and across evidence frame durations from 16.7ms to 167ms. This underlines that it is a robust empirical phenomenon that is important to explain for any theory of perceptual decision-making. Second, we find that the results from our first experiment clearly favor an algorithmic explanation of the primacy effect: the strength of the primacy effect agrees between conditions as a function of stimulus frame (Fig. 2A,B top) much better than as a function of physical time (Fig. 2A,B bottom). Third, these results are confirmed by our second experiment for frame durations of 42ms and 83ms. For the shortest frame duration of 17ms, the results are more complicated: better agreeing with the 42ms weight profile when aligned by physical time which may be explained by the temporal extend of neural receptive fields (a biophysical feature) playing a larger role in the integration of evidence across multiple evidence frames.

Conclusions: Overall, our results suggest that the primacy effect is best understood as a suboptimality arising on the algorithmic level describing how the brain accumulates evidence over time, and not as the consequence of biophysical constraints, e.g. how evidence accumulation is implemented by attractor dynamics within or across cortical areas. While algorithmic models, too, are of course implemented as a biophysical system in the brain, our results demonstrate that the corresponding implementation adapts to the time scale present in the inputs, e.g. increasing integration time constants as needed to account for longer evidence frames. Finally, our finding of a robust primacy effect even on the scale of a few 100ms indicates that this suboptimality is an integral part of perceptual decision-making on time scales relevant for natural behavior, and not a side-effect of artificially prolonged stimulus durations in the lab.

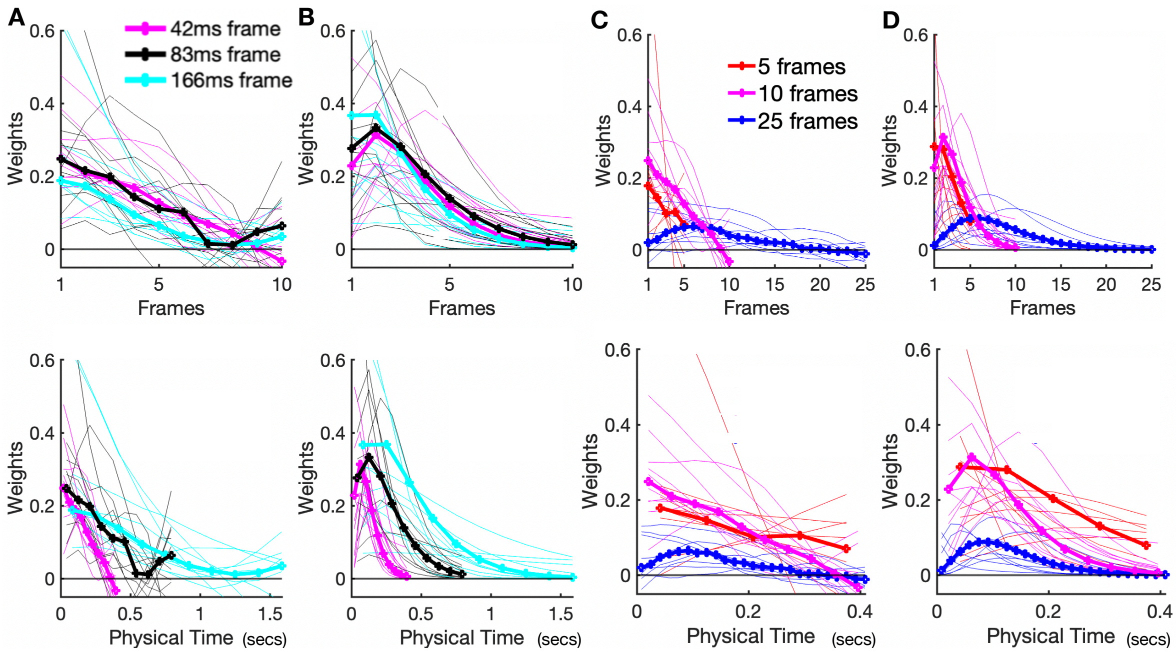

Figure 2: **A-B:** Compares data in experiments (1) (2) (3) **A:** Logistic regression based PK compared for trials with fixed number of stimulus frames (top: plotted against stimulus frame number, bottom: plotted against physical time in a trial **B:** Gamma pdf fit PKs (top and bottom same as **A**) **C-D:** Same as **A-B** but compares data in experiments (2) (4) (5)

References

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological review*, *113*(4), 700.

Drugowitsch, J., Wyart, V., Devauchelle, A.-D., & Koechlin, E. (2016). Computational Precision of Mental Inference as Critical Source of Human Choice Suboptimality. *Neuron*, *92*(6), 1398–1411.

Glaze, C. M., Filipowicz, A. L., Kable, J. W., Balasubramanian, V., & Gold, J. I. (2018). A bias–variance trade-off governs individual differences in on-line learning in an unpredictable environment. *Nature Human Behaviour*, *2*(3), 213–224.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual review of neuroscience*, *30*(30), 535–574.

Kiani, R., Hanks, T. D., & Shadlen, M. N. (2008). Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *The Journal of neuroscience*, *28*(12), 3017–3029.

Lange, R. D., Chattoraj, A., Beck, J. M., Yates, J. L., & Haefner, R. M. (2021). A confirmation bias in perceptual decision-making due to hierarchical approximate inference. *PLOS Computational Biology*, *17*(11), e1009517.

Nienborg, H., & Cumming, B. G. (2009). Decision-related activity in sensory neurons reflects more than a neuron's causal effect. *Nature*, *459*(7243), 89–92.

Wang, X.-J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron*, *36*(5), 955–968.

Yates, J. L., Park, I. M., Katz, L. N., Pillow, J. W., & Huk, A. C. (2017). Functional dissection of signal and noise in mt and lip during decision-making. *Nature neuroscience*, *20*(9), 1285–1292.