Multisensory Context Warps Time Perception

Lucy Lai (lucy.lai@rice.edu)
Department of Cognitive Sciences
Rice University
6100 Main St, Houston, TX 77005 USA

John F. Magnotti (john.magnotti@bcm.edu)
Department of Neurosurgery
Baylor College of Medicine
1 Baylor Plaza, Houston, TX 77030 USA

Jeffrey M. Yau (jeffrey.yau@bcm.edu)
Department of Neuroscience
Baylor College of Medicine
1 Baylor Plaza, Houston, TX 77030 USA

Abstract

How the brain processes timing information from multiple senses remains poorly understood. Here, we explore how audition and touch interact in the time domain. We tested how the duration of ignored sounds influences the perceived duration of co-occurring tactile events. Distractor sounds exerted systematic attractive and repulsive biasing effects on tactile duration judgments that were consistent across multiple timing ranges. We developed a two-step observer model to explain these results. First, the observer decides to bind or to separate the auditory and tactile duration cues using causal inference. Subsequently, the observer computes a Bayesian estimate of duration using either a coupling or decoupling prior depending on the decision to bind or to separate cues, respectively. While existing cue combination models comparably predict attractive perceptual biases, the two-step model, owing to its conditional decoupling prior, is the only model that also accounts for large repulsive biases. Critically, the model predicts that increased sensory uncertainty shifts repulsive biases toward attraction, which we validated in separate experiments. These results imply multisensory computations are conditioned on probabilistic decisions to bind or to separate sensory cues. Our model provides a unified framework for understanding the extensive and flexible perceptual outcomes that result from multisensory cue interactions.

Keywords: multisensory; time perception; Bayesian inference; causal inference; psychophysics; computational modelling

Background

Complex human behavior, such as motor coordination and speech perception, relies on the ability to accurately perceive the timing of events. Yet perceived durations are malleable and influenced by conflicting information as well as the context in which they are presented. Despite the fundamental importance of time perception, we still lack a theoretical framework for understanding how humans process timing information from multiple senses. Because audition and touch share mechanisms for frequency processing (Yau, Olenczak, Dammann, & Bensmaia, 2009), we leverage this relationship to better understand how the brain binds or separates multisensory time information. In the present study, we use human psychophysics to understand how the duration of sounds influences the perceived duration of co-occurring tactile events. Furthermore, we develop a two-step observer model that can describe the wide range of perceptual outcomes that result from multisensory duration interactions.

Methods

To quantify tactile duration processing, we asked human subjects to perform a two-alternative forced choice (2AFC) task in the presence and absence of auditory distractors. Subjects (N=16) felt two sequential vibrations delivered to a finger and reported which of the two stimuli appeared longer in duration. Each trial contained a standard 500-ms stimulus and a comparison stimulus that varied from 300 to 700ms.

To establish the influence of sound duration on tactile duration judgements, we paired the standard stimulus with five distractor sounds whose durations ranged from 300 to 700ms. Subjects performed the tactile duration discrimination task while explicitly ignoring the sounds.

Auditory distractors exert attractive and repulsive biases on tactile duration perception

We observed robust and systematic bias effects that depended on the relative durations of the sounds and vibrations (*Figure 1b*). Shorter sounds compressed the perceived duration of a tactile stimulus, causing it to appear shorter (e.g., a 300-ms sound causes a 500-ms tactile event to be perceived as $\sim$450ms). Counterintuitively, longer sounds also caused the tactile stimulus to appear shorter (e.g., a 700-ms sound causes a 500-ms tactile event to be perceived as $\sim$450ms). In other words, shorter distractors biased percieved tactile duration *towards* the auditory distractor duration (attraction), while longer distractors biased percieved tactile duration *away* from the distractor duration (repulsion).

This surprising result shows that multisensory timing cues, depending on their relative durations, can lead to attractive or repulsive perceptual interactions. This novel pattern, which generalized over a range of tested durations, implies that the nervous system accounts for more than just timing cues when computing perceived duration in a multisensory context.

A Bayesian model of context-dependent multisensory time perception

We developed a probabilistic observer model to explain our observations quantitatively and to provide a framework for understanding the computations that yield attractive and repulsive biases. We used the observer model to simulate performance on the tactile discrimination task. On any given trial interval, the model produces a duration estimate in a two-step process. First, the observer decides to bind or to separate the auditory and tactile duration cues using causal inference. Subsequently, the observer computes a Bayesian estimate of duration using either a coupling or decoupling prior depending the initial decision to bind or to separate, respectively.

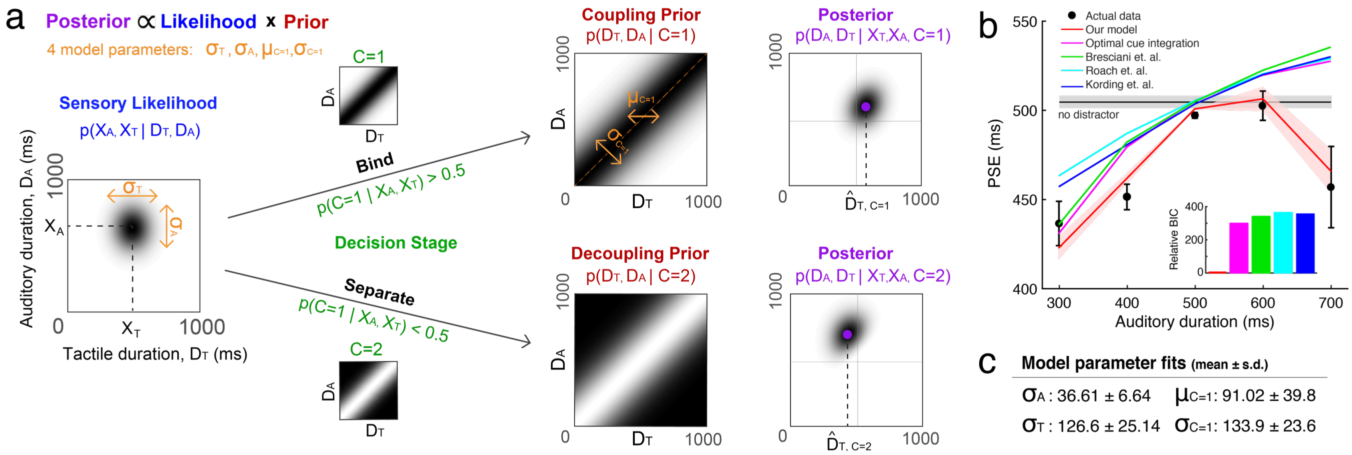

Figure 1: **(a)** Two-step Bayesian inference model. **(b)** Point of subjective equality as a function of auditory distractor duration. Colored lines indicate model predictions. (Inset) Relative BIC for the models. **(c)** Best fitting model parameters.

Figure 1a provides an illustration of the model architecture for an example trial interval comprising a vibration of duration D_T and a sound of duration D_A. Due to sensory noise, the likelihood is represented by a bivariate Gaussian distribution centered at (X_T, X_A). First, the model observer must decide whether to bind or to separate the two sensory cues through causal inference (Kording et al., 2007). Concretely, the observer calculates $p(C = 1|X_T, X_A)$, or the probability that the measured cues came from one cause. If $p(C = 1|X_T, X_A) \geq p(C = 2|X_T, X_A)$, the observer commits to cue binding. In all other cases, the observer commits to cue separation. Critically, this initial decision conditions subsequent operations (Stocker & Simoncelli, 2008).

If the observer decides to bind cues, it multiplies the likelihood, $p(X_T, X_A)$, by a coupling prior that represents the joint distribution of durations from the two sensory channels that are bound, $p(D_T, D_A|C = 1)$. Alternatively, if the observer decides to separate cues, it multiplies the likelihood by a *de*coupling prior that represents the joint distribution of durations from the two channels that are separated, $p(D_T, D_A|C = 2)$. The coupling and decoupling prior are related by $p(D_A, D_T|C = 2) + p(D_A, D_T|C = 1) = 1$. The posterior, or the perceived tactile duration, is a combination of the likelihood and the conditional prior.

The model is fully described by four parameters: $\sigma_A, \sigma_T, \sigma_{C=1}$, and $\mu_{C=1}$. The best fitting parameters (*Figure 1c*) reveal that auditory sensitivities are more reliable than tactile sensitivities, $\sigma_A \leq \sigma_T$, and that there exists moderate coupling of audio-tactile cues ($\sigma_{C=1} > 0$). Additionally, $\mu_{C=1} \geq 0$, revealing that observers more readily infer a single event when tactile durations are longer than auditory durations. These four parameters can be described as constants of proportionality (e.g. $W_T = \sigma_T/D_T$) allowing generalization to arbitrary duration ranges.

We ran simulations to obtain model predictions for each distractor condition and found good agreement with the behavioral data (*Figure 1b, red*). While existing models of multisensory cue combination predict perceptual outcomes ranging from cue independence to forced fusion (Ernst & Banks, 2002; Roach, Heron, & McGraw, 2006; Kording et al., 2007), the two-step model, is the only model that also accounts for the robust repulsive biases in the behavioral data (*Figure 2b; effect of 700-ms distractor*). Critically, the model predicts that increased sensory uncertainty shifts repulsive biases toward attraction, which we validated in additional behavioral experiments.

Our work reveals novel ways in which multisensory context causes attractive and repulsive interactions in time perception and provides a unified theoretical framework for understanding the extensive and flexible perceptual outcomes that result from multisensory cue interactions.

Acknowledgments

LL and JMY conceived and designed the study. LL performed the experiments. LL and JMY analyzed the data and interpreted results. LL, JFM, and JMY and developed the model. This work was supported by a Sloan Research Fellowship and an NIH-R01NS097462 to JMY.

References

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429433. doi: 10.1038/415429a

Kording, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal inference in multisensory perception. *PLoS ONE*, *2*(9). doi: 10.1371/journal.pone.0000943

Roach, N. W., Heron, J., & McGraw, P. V. (2006, Jul). Resolving multisensory conflict: a strategy for balancing the costs and benefits of audio-visual integration. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1598), 21592168. doi: 10.1098/rspb.2006.3578

Stocker, A. A., & Simoncelli, E. P. (2008). A bayesian model of conditioned perception. In J. C. Platt, D. Koller, Y. Singer, & S. T. Roweis (Eds.), *Advances in neural information processing systems 20* (pp. 1409–1416).

Yau, J. M., Olenczak, J. B., Dammann, J. F., & Bensmaia, S. J. (2009). Temporal frequency channels are linked across audition and touch. *Current Biology*, *19*(7), 561566. doi: 10.1016/j.cub.2009.02.013