Choice history biases evidence accumulation: a cross-species comparison from humans to mice

Kianté Fernandez[1], **Alexander Fengler**[2], **Anne E. Urai**[3,*]

***For correspondence:** a.e.urai@fsw.leidenuniv.nl (AEU)

[1]University of California, Los Angeles, United States; [2]Brown University, United States; [3]Cognitive Psychology Unit, Leiden University, Netherlands

Abstract Mice are increasingly used to study the neural circuit-level basis of behavior, often with the ultimate goal of extrapolating these insights to humans. To generalize insights about neural functioning across species, it is crucial to ensure correspondence in behavioral and cognitive strategy. We previously showed that human observers' evidence accumulation is biased by previous choices (***Urai et al., 2019***). To replicate these findings across species, we fit Diffusion Decision Models (***Fengler et al., 2025***) to behavioral data from 62 mice performing a standardized perceptual decision-making task (***The International Brain Laboratory et al., 2021***). We identified the same cognitive strategy of history-dependent evidence accumulation: individual differences in choice repetition were explained by a history-dependent bias in the rate of evidence accumulation rather than its starting point. We argue that history-biased evidence integration reflects a fundamental aspect of perceptual decision-making, that may transcend the specific species.

Introduction

In perceptual decision-making, evidence accumulation is biased by previous choices: our previous work has shown that in human this seems to happen through a history-dependent change in the speed of evidence accumulation (drift bias), rather than in via a shift starting point of the accumulation process (***Urai et al., 2019***) (Figure 1). Choice history (defined as one or more past choices) thus seems to bias the interpretation of current sensory input, akin to shifting endogenous attention toward (or away from) the previously selected interpretation of sensory evidence (***Urai and Donner, 2022***). This decision-making strategy, which robustly captures individual differences across tasks, may be exhibited across mammalian species—thereby providing insights into the shared neural circuit mechanisms of cognition.

Recent advances in training mice to perform complex tasks, combined with powerful neural measurement tools, have positioned mice as a popular model species in cognitive neuroscience research (***Urai et al., 2022***). An important premise is that cognitive and computational mechanisms of behavior are preserved across species. However, this is rarely explicitly tested, creating challenges in the ability to translate mice-based neuroscience findings to humans (***Barron et al., 2020***). Here, we directly replicate our previous analyses in humans (***Urai et al., 2019***) in mice. We analyzed choices and response times of 62 mice performing a standardized decision task (***The International Brain Laboratory et al., 2021***) and show that animals exhibit the same computational strategy as humans: a history-dependent drift bias.

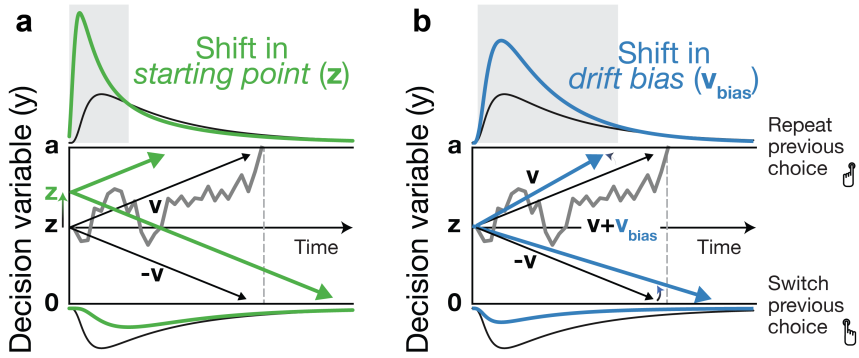

Figure 1. The DDM postulates that noisy sensory evidence is accumulated over time, until the resulting decision variable y reaches one of two bounds (lines at $y = 0$ and $y = a$) for the two choice options. Repeating this process over many trials yields RT distributions for each choice option respectively plotted above and below the bounds. Gray line: example trajectory of the decision variable from a single trial. Black lines: mean drift and resulting RT distributions under unbiased conditions. **(a)** Choice history-dependent shift in starting point. Green lines: mean drift and RT distributions under biased starting point. Gray-shaded area indicates those RTs for which starting point leads to choice bias. **(b)** Choice history-dependent shift in drift bias. Blue lines: mean drift and RT distributions under biased drift. Gray shaded area indicates those RTs for which drift bias leads to choice bias. Adapted under a CC-BY license from (***Urai et al., 2019***).

Results

Choice repetition in perceptual tasks has been reported across species from rodents to monkeys and humans (***Busse et al., 2011***; ***Gold et al., 2008***; ***Fründ et al., 2014***). The basis of this work is an open dataset from the International Brain Laboratory (***The International Brain Laboratory et al., 2021***), where animals are trained to detect the location of a visual stimulus on a screen, which varies in contrast (Figure 2a). Animals then use a wheel to move the stimulus into the center of the screen to obtain a water reward.

We here selected data from early in the animals' training: after animals had sufficiently learned the basic task, but before the introduction of a block manipulation that complicates analyses of choice repetition. As shown in previous analyses of these data (***The International Brain Laboratory et al., 2021***; ***Findling et al., 2025***), mice tend to repeat their past choice irrespective of the past trial's outcome (correct or error) (Figure 2).

Fitting history-dependent sequential sampling models

To distinguish different mechanisms that may underlie this repetition bias, we fit four variants of a Diffusion-Decision Model (***Ratcliff et al., 2016***): with choice history affecting no parameters, only z, only v_{bias}, or both z and v_{bias}. Model comparison favors models where only v_{bias} or both parameters scale with choice history, replicating our previous results in humans (Figure 3a).

The results from quantitative model comparisons were corroborated by inspection of individual differences. Animals' choice repetition probability correlated with their history shift in v_{bias}, but not in z (Figure 3b). Conditional bias functions confirm this qualitatively: animals show repetition bias across all RT quantiles, a pattern expected from a v_{bias} but not z (Figure 3). This replicates our previous finding in six human perceptual decision-making datasets (***Urai et al., 2019***).

Discussion

In this work, we test various formulations of sequential sampling models (***Fengler et al., 2022***) to a public dataset of mouse decision-making (***The International Brain Laboratory et al., 2021***) and find that mice display the same history-dependent drift biases as previously shown in humans (***Urai et al., 2019***). Evidence accumulation over multiple temporal scales may thus reflect a fundamental aspect of decision-making, conserved across mammalian species.

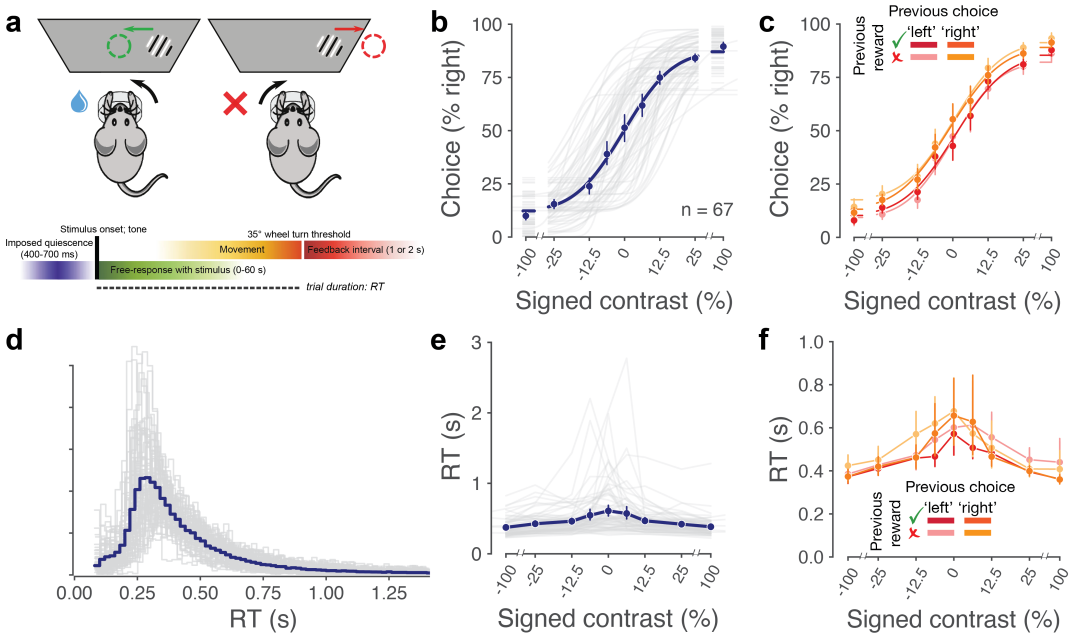

Figure 2. Mice repeat their choices. **(a)** A standardized decision-making task, adapted from (***The International Brain Laboratory et al., 2021***) under a CC-BY license. **(b)** Psychometric functions for all animals (each shown with a grey line). **(c)** History-conditioned psychometric functions. **(d)** Distributions of response times, taken as the time from stimulus onset until response completion: either bringing the stimulus into the center of the screen, or the equivalent distance off the screen. Each animal is shown with one grey line. **(e)** Chronometric functions, showing median RT per contrast level for all animals (each shown with a grey line). **(f)** History-conditioned chronometric functions.

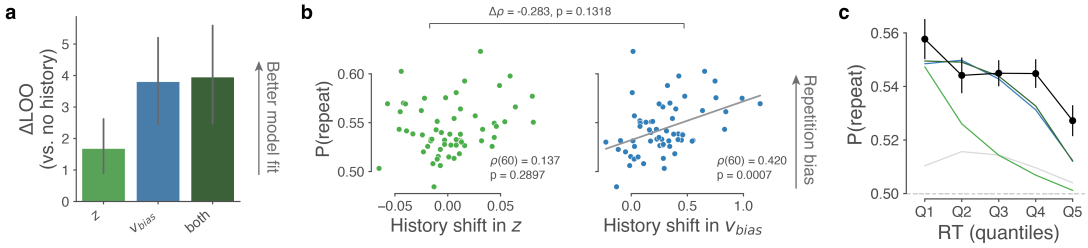

Figure 3. Individual choice history biases are explained by history-dependent changes in drift bias, not starting point. **(a)**. Model comparison of models where z, v_{bias} or both were allowed to vary as a function of the animals' previous choice. The LOO for a model without history dependence was used as a baseline for each dataset. Higher ΔLOO values indicate a model that is better able to explain the data, taking into account the model complexity. **(b)** Correlation between individual choice repetition probabilities, P(repeat), and history shift in starting point (left column, green) and drift bias (right column, blue). Parameter estimates were obtained from a model in which both bias terms were allowed to vary with previous choice. Grey line shows the best fit of a linear regression, only plotted for Spearman's ρ correlations significant at p<0.05. Top: $\Delta\rho$ value between the two correlation coefficients, with p-value computed from a Steiger's test. **(c)** Conditional bias functions comparing observed data with simulations from four DDM variants. The probability of repeating a previous choice, P(repeat), is plotted across five reaction time (RT) quantiles. The observed data (black) show that choice repetition is most frequent on fast trials and decreases as RT increases. This empirical pattern is captured by the best-fitting v_{bias} model (blue).

These findings set the stage for linking the computations of decision-making to neural dynamics at the single-cell and population levels. Specifically, recent work showed that an exponentially weighted average of past choices informs an animal's prior, which can be decoded across brain regions (***Findling et al., 2025***). Future work could extend these findings by developing neurocognitive models that link neural activity at both the single-cell and population levels with behavioral modalities through integrative modeling (***Nunez et al., 2024***). Such joint modeling can enable the simultaneous characterization of single-trial neural measures and cognitive modeling parameters, thereby capturing trial-by-trial variations in drift bias (***Ghaderi-Kangavari et al., 2023***).

Our replication of previous human findings has several limitations that we plan to address in future work. We used animals' single past choice at $t-1$ as a limited read-out for the full prior that is constructed from past evidence: future work could include choice history beyond one trial (***Fründ et al., 2014***; ***Findling et al., 2025***; ***Urai et al., 2019***). Moreover, choice history bias may arise both from true repetition behavior and from slow drifts in decision criterion (***Gupta and Brody, 2022***; ***Vloeberghs et al., 2025***; ***Urai, 2025***), which may in turn affect trial-to-trial fluctuations in drift bias. Lastly, mice are known to occupy different engagement states over the course of a session (***Ashwood et al., 2022***; ***Urai et al., 2019***), which may need to be estimated to better fit sequential sampling models (***Chakravarty et al., 2025***). All of these extensions may ultimately allow a richer estimation of the interaction between latent computational variables that vary across and within trials (***Urai, 2025***).

Methods and Materials

Task, procedure and data selection

Mouse data were collected by the International Brain Laboratory (***The International Brain Laboratory et al., 2021***), and released as part of (***Bruijns et al., 2023***). Mice were trained to perform a perceptual decision-making task, reporting the location of a visual grating (of varying contrast) by turning a steering wheel left or right. We selected sessions where mice had mastered the basic task but had not yet been exposed to structured stimulus sequences with biased probabilities. This approach ensured that behavioral choice history biases reflected endogenous, rather than experimentally induced, biases. For each mouse, we analyzed three sessions just prior to achieving training proficiency in the basic task, ensuring comparable performance levels across animals while avoiding any influence of the probabilistic stimulus sequences introduced in later training stages.

We defined RT as the difference between stimulus onset and completion of the response. Note that this differs slightly from previous work on this dataset (***The International Brain Laboratory et al., 2021***), which used the time of first movement initiation rather than choice completion as reaction time. While likely a mixture of multiple processes, we chose to include the entire interval as our measure of the decision process, as we aim to evaluate the DDMs capacity to capture components of processing associated with the entire period the animal engaged in the task.

We included sessions based on several quality control criteria: each session included at least five contrast levels; performance above 80% correct on high-contrast trials (≥50% contrast); a median RT of less than 1 second; no negative RTs (likely reflecting hardware failure); and a total of >= 200 trials per session remaining after these exclusion criteria. We included only trials with RTs and movement onset times between 0.08 and 2 seconds - excluding outliers that likely represented anticipatory or disengaged responses (***The International Brain Laboratory et al., 2025***). After exclusions, our final dataset contained choices and RTs from 62 animals.

Cognitive Modeling

To model the choice and RTs of the animals, we use the sequential sampling modeling (SSM) framework (***Forstmann et al., 2016***; ***Fengler et al., 2022***). This widely used class of cognitive models describes decision-making as a stochastic, dynamic evidence accumulation process which evolves

one or multiple momentary evidence-states, until a particular predetermined threshold level is reached which triggers a given choice (***Ratcliff and McKoon, 2008***; ***Ratcliff, 1978***). Such "boundary-crossings" trigger a response, leading to distributions of RTs. SSMs thus provide a generative model for reaction time distributions of actions made by organisms, capturing both the speed of the response and the choice itself. Since similar models were used to capture choice history effects in previous work, we only briefly reintroduce the models here and refer readers to ***Urai et al.*** (***2019***) for further details.

DDM

In this study we use the basic Diffusion-Decision Model (***Ratcliff et al., 2016***) (also known as Drift Diffusion Model; Figure 1).

The DDM describes the accumulation of noisy sensory evidence which can be described via the following stochastic differential equation:

$$dx = v \cdot dt + c\mathrm{d}W$$

where x is the decision variable and cdW is Gaussian distributed white noise with mean 0 and variance c^2dt. This represents a basic continuous random walk with drift (***Smith, 2000***). The evidence accumulation process terminates when x reaches either the upper bound a or the lower bound 0 (sometimes this is instead represented as hitting a or $-a$ instead).

In our work, we represent the drift v_t, where t refers to a given trial number, via the following simple linear regression:

$$v_t = \beta^v_{intercept} + \beta^v_{contrast} \cdot s_t + \beta^v_{bias} \cdot h_t$$

where s_t represents the signed stimulus contrast on trial t (negative for left stimuli, positive for right stimuli). Since we observed a saturation of drift rates for contrasts higher than 25%, we clipped all contrast levels above this value to 25% (or -25% for stimuli on the left). h_t is defined as the signed choice (-1 for leftwards, +1 for rightwards) on trial $t-1$.

We use a similar definition for the starting point:

$$x(0)_t = z_t = \beta^z_{intercept} + \beta^z_{bias} \cdot h_t$$

where β^z_{bias} represents how previous choices affect the starting point—a history-dependent bias in the decision process. We compare these against models without history-dependence where we either drop the regression aspect on v and/or the regression aspect on z, in which case they act as parameters without trial-level adjustments.

Together, β^v_{bias} and β^z_{bias} capture how choice history influences decisions through distinct mechanisms. As previously shown, the two biasing mechanisms result in the same (asymmetric) fraction of choices, but they differ in terms of the resulting shapes of RT distributions (Figure 1).

Model comparison

To compare model fits across model variants, we estimated out-of-sample predictive performance using LOO cross-validation (***Vehtari et al., 2017***). This approximation, based on Pareto-smoothed importance sampling (PSIS), uses the expected log pointwise predictive density (ELPD) estimates to assess how well a model is likely to fit holdout data without requiring iterative refitting (***Vehtari et al., 2017***). We used the routines implemented in the ArviZ Python package (***Kumar et al., 2019***).

Parameter estimation

All models were fit using HSSM (Hierarchical Sequential Sampling Modeling) (***Fengler et al., 2025***), a Python toolbox that facilitates Bayesian inference for a broad class of Sequential Sampling Models, including hierarchical mixed-effects regressions on each underlying cognitive model parameter. To do so, HSSM constructs specified regressions with the Bambi library (***Capretto et al., 2022***) and complies the models with the PyMC probabilistic programming library (***Abril-Pla et al., 2023***). Parameter posteriors and posterior predictive checks were plotted with ArviZ (***Kumar et al., 2019***).

Data and code availability

All code is available at https://github.com/kiante-fernandez/2023_choicehistory_HSSM.

Acknowledgments

We thank the International Brain Laboratory for data collection, curation and software development. Philippa Johnson, Charline Tessereau, Naureen Ghani, Naoki Hiratani, Miles Wells, Peter Latham and Alex Pouget provided valuable discussions on the definition of RT in the IBL task. We thank Zeynep Gunes Ozkan for discussions and help in an earlier stage of the project. AEU is supported by a Veni grant (VI.Veni.212.184) from the Netherlands Organization for Scientific Research.

References

Abril-Pla O, Andreani V, Carroll C, Dong L, Fonnesbeck CJ, Kochurov M, Kumar R, Lao J, Luhmann CC, Martin OA, Osthege M, Vieira R, Wiecki T, Zinkov R. PyMC: a modern, and comprehensive probabilistic programming framework in Python. PeerJ Computer Science. 2023; 9:e1516. doi: 10.7717/peerj-cs.1516.

Ashwood ZC, Roy NA, Stone IR, International Brain Laboratory, Urai AE, Churchland AK, Pouget A, Pillow JW. Mice alternate between discrete strategies during perceptual decision-making. Nature Neuroscience. 2022 Feb; 25(2):201–212. doi: 10.1038/s41593-021-01007-z.

Barron HC, Mars RB, Dupret D, Lerch JP, Sampaio-Baptista C. Cross-species neuroscience: closing the explanatory gap. Philosophical Transactions of the Royal Society B: Biological Sciences. 2020 Nov; 376(1815):20190633. https://royalsocietypublishing.org/doi/10.1098/rstb.2019.0633, doi: 10.1098/rstb.2019.0633, publisher: Royal Society.

Bruijns SA, Laboratory IB, Bougrova K, Laranjeira IC, Lau PYP, Meijer GT, Miska NJ, Noel JP, Pan-Vazquez A, Roth N, Socha KZ, Urai AE, Dayan P, Dissecting the Complexities of Learning With Infinite Hidden Markov Models. bioRxiv; 2023. https://www.biorxiv.org/content/10.1101/2023.12.22.573001v1, doi: 10.1101/2023.12.22.573001, pages: 2023.12.22.573001 Section: New Results.

Busse L, Ayaz A, Dhruv NT, Katzner S, Saleem AB, Schölvinck ML, Zaharia AD, Carandini M. The Detection of Visual Contrast in the Behaving Mouse. Journal of Neuroscience. 2011 Aug; 31(31):11351–11361. http://www.jneurosci.org/content/31/31/11351, doi: 10.1523/JNEUROSCI.6689-10.2011.

Capretto T, Piho C, Kumar R, Westfall J, Yarkoni T, Martin OA. Bambi: A Simple Interface for Fitting Bayesian Linear Models in Python. Journal of Statistical Software. 2022; 103(15):1–29. https://www.jstatsoft.org/index.php/jss/article/view/v103i15, doi: 10.18637/jss.v103.i15.

Chakravarty S, Delgado-Sallent C, Kane GA, Xia H, Do QH, Senne RA, Scott BB. A cross-species framework for investigating perceptual evidence accumulation. eLife. 2025 Apr; 14. https://elifesciences.org/reviewed-preprints/105116, doi: 10.7554/eLife.105116.1, publisher: eLife Sciences Publications Limited.

Fengler A, Xu Y, Bera K, Omar A, Frank M. HSSM: A generalized toolbox for hierarchical Bayesian estimation of computational models in cognitive neuroscience. Fengler, A, Xu, P, Bera, K, Omar, A, Frank, MJ (in preparation). 2025; .

Fengler A, Bera K, Pedersen ML, Frank MJ. Beyond drift diffusion models: Fitting a broad class of decision and reinforcement learning models with HDDM. Journal of cognitive neuroscience. 2022; 34(10):1780–1805. https://direct.mit.edu/jocn/article-abstract/34/10/1780/112585, publisher: MIT Press One Broadway, 12th Floor, Cambridge, Massachusetts 02142, USA ....

Findling C, Hubert F, Laboratory IB, Acerbi L, Benson B, Benson J, Birman D, Bonacchi N, Bruijns S, Carandini M, Catarino JA, Chapuis GA, Churchland AK, Dan Y, Davatolhagh F, DeWitt EE, Engel TA, Fabbri M, Faulkner M, Fiete IR, et al. Brain-wide representations of prior information in mouse decision-making. Nature. 2025; (in press). https://www.biorxiv.org/content/10.1101/2023.07.04.547684v5, doi: 10.1101/2023.07.04.547684, pages: 2023.07.04.547684 Section: New Results.

Forstmann BU, Ratcliff R, Wagenmakers EJ. Sequential Sampling Models in Cognitive Neuroscience: Advantages, Applications, and Extensions. Annual Review of Psychology. 2016; 67(1):641–666. https://doi.org/10.1146/annurev-psych-122414-033645, doi: 10.1146/annurev-psych-122414-033645, _eprint: https://doi.org/10.1146/annurev-psych-122414-033645.

Fründ I, Wichmann FA, Macke JH. Quantifying the effect of intertrial dependence on perceptual decisions. Journal of Vision. 2014 Jun; 14(7):9–9. http://www.journalofvision.org/lookup/doi/10.1167/14.7.9, doi: 10.1167/14.7.9.

Ghaderi-Kangavari A, Rad JA, Nunez MD. A General Integrative Neurocognitive Modeling Framework to Jointly Describe EEG and Decision-making on Single Trials. Computational Brain & Behavior. 2023 Sep; 6(3):317–376. https://doi.org/10.1007/s42113-023-00167-4, doi: 10.1007/s42113-023-00167-4.

Gold JI, Law CT, Connolly P, Bennur S. The Relative Influences of Priors and Sensory Evidence on an Oculomotor Decision Variable During Perceptual Learning. Journal of Neurophysiology. 2008 Aug; 100(5):2653–2668. http://jn.physiology.org/cgi/doi/10.1152/jn.90629.2008, doi: 10.1152/jn.90629.2008.

Gupta D, Brody CD. Limitations of a proposed correction for slow drifts in decision criterion. Neurons, Behavior, Data analysis, and Theory. 2022 May; 1:1–19. https://nbdt.scholasticahq.com/article/35908-limitations-of-a-proposed-correction-for-slow-drifts-in-decision-criterion, doi: 10.51628/001c.35908, publisher: The neurons, behavior, data analysis and theory collective.

Kumar R, Carroll C, Hartikainen A, Martin O. ArviZ a unified library for exploratory analysis of Bayesian models in Python. Journal of Open Source Software. 2019; 4(33):1143. https://doi.org/10.21105/joss.01143, doi: 10.21105/joss.01143, publisher: The Open Journal.

Nunez MD, Fernandez K, Srinivasan R, Vandekerckhove J. A tutorial on fitting joint models of M/EEG and behavior to understand cognition. Behavior Research Methods. 2024 Feb; https://doi.org/10.3758/s13428-023-02331-x, doi: 10.3758/s13428-023-02331-x.

Ratcliff R. A theory of memory retrieval. Psychological Review. 1978; 85:59–108. https://semanticscholar.org/paper/bccf38995692eae3f555e91093e31f9b0199d040, doi: 10.1037/0033-295X.85.2.59, tex.ids= ratcliffTheoryMemoryRetrieval1978.

Ratcliff R, McKoon G. The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. Neural computation. 2008 Apr; 20(4):873–922. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2474742/, doi: 10.1162/neco.2008.12-06-420.

Ratcliff R, Smith PL, Brown SD, McKoon G. Diffusion Decision Model: Current Issues and History. Trends in Cognitive Sciences. 2016 Apr; 20(4):260–281. https://www.sciencedirect.com/science/article/pii/S1364661316000255, doi: 10.1016/j.tics.2016.01.007, tex.ids= ratcliffDiffusionDecisionModel2016a place: Netherlands publisher: Elsevier Science.

Smith PL. Stochastic dynamic models of response time and accuracy: A foundational primer. Journal of mathematical psychology. 2000; 44(3):408–463. https://www.sciencedirect.com/science/article/pii/S0022249699912609, publisher: Elsevier.

The International Brain Laboratory, Aguillon-Rodriguez V, Angelaki D, Bayer H, Bonacchi N, Carandini M, Cazettes F, Chapuis G, Churchland AK, Dan Y, Dewitt E, Faulkner M, Forrest H, Haetzel L, Häusser M, Hofer SB, Hu F, Khanal A, Krasniak C, Laranjeira I, et al. Standardized and reproducible measurement of decision-making in mice. eLife. 2021 May; 10:e63711. https://doi.org/10.7554/eLife.63711, doi: 10.7554/eLife.63711, publisher: eLife Sciences Publications, Ltd.

The International Brain Laboratory, Benson B, Benson J, Birman D, Bonacchi N, Carandini M, Catarino JA, Chapuis GA, Churchland AK, Dan Y, Dayan P, DeWitt EE, Engel TA, Fabbri M, Faulkner M, Fiete IR, Findling C, Freitas-Silva L, Gercek B, Harris KD, et al. A Brain-Wide Map of Neural Activity during Complex Behaviour. Nature. 2025; (in press). https://www.biorxiv.org/content/10.1101/2023.07.04.547681v2, doi: 10.1101/2023.07.04.547681, section: New Results.

Urai AE. Structure uncovered: understanding temporal variability in perceptual decision-making. Trends in Cognitive Sciences. 2025 Jul; 0(0). https://www.cell.com/trends/cognitive-sciences/abstract/S1364-6613(25)00149-4, doi: 10.1016/j.tics.2025.06.003, publisher: Elsevier.

Urai AE, Doiron B, Leifer AM, Churchland AK. Large-scale neural recordings call for new insights to link brain and behavior. Nature Neuroscience. 2022 Jan; 25(1):11–19. https://www.nature.com/articles/s41593-021-00980-9, doi: 10.1038/s41593-021-00980-9, publisher: Nature Publishing Group.

Urai AE, Donner TH. Persistent activity in human parietal cortex mediates perceptual choice repetition bias. Nature Communications. 2022 Oct; 13(1):6015. https://www.nature.com/articles/s41467-022-33237-5, doi: 10.1038/s41467-022-33237-5, publisher: Nature Publishing Group.

Urai AE, de Gee JW, Tsetsos K, Donner TH. Choice history biases subsequent evidence accumulation. eLife. 2019 Jul; 8:e46331. https://doi.org/10.7554/eLife.46331, doi: 10.7554/eLife.46331, publisher: eLife Sciences Publications, Ltd.

Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Statistics and Computing. 2017 Sep; 27(5):1413–1432. https://doi.org/10.1007/s11222-016-9696-4, doi: 10.1007/s11222-016-9696-4.

Vloeberghs R, Urai AE, Desender K, Linderman SW. A Bayesian hierarchical model of trial-to-trial fluctuations in decision criterion. PLOS Computational Biology. 2025 Jul; 21(7):e1013291. https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1013291, doi: 10.1371/journal.pcbi.1013291, publisher: Public Library of Science.