Joint Modeling of Choices and Response Times in Multi-stage Decisions via Likelihood Approximation

Jialin Li (jl16057@nyu.edu)[1], Carlos G. Correa (carlos.correa@nyu.edu)[1], Doris Yu (dorisyu@uw.edu)[3], Prakhar Godara (pg2991@nyu.edu)[1] & Marcelo G. Mattar (mm13100@nyu.edu)[1,2]

[1]Department of Psychology, New York University
[2]Center of Neural Science, New York University
[3]Foster School of Business, University of Washington

Abstract

Planning involves a process of considering future states before acting. To understand this process, researchers typically infer planning algorithms by fitting computational models to choices. However, different planning models often predict the same choices, despite relying on different computations. Reaction time can help distinguish among models, since different computations produce different temporal signatures. However, incorporating reaction time into fitting is challenging because analytical likelihoods are typically unavailable. Here we propose a likelihood-free method to estimate the density for choices and reaction times in multi-stage decision making. We validate the method through comparisons with analytical solutions, parameter recovery, and showing robust estimates relative to distribution-free and summary statistic approaches. Through a new human experiment and fitting evidence accumulation models from Solway and Botvinick (2015), we demonstrate that modeling the full distribution is important to explain human behavior. Overall, our method is a valuable tool for modeling reaction times in multi-stage decision-making.

Keywords: planning; evidence accumulation; decision tree; likelihood free estimation

Introduction

Every day, people engage in sequential decisions that require mentally simulating multiple future outcomes before acting. This ability to anticipate and evaluate future consequences is a defining characteristic of planning (Mattar and Lengyel, 2022), formalized as a tree search problem in which people construct and search a decision tree that links present actions with distant rewards (Kuperwajs et al., 2025).

Because planning is an internal, unobservable process, researchers must infer its structure indirectly from behavior. Extensive work has modeled human planning by fitting computational models to choices in multi-stage decision-making (Daw et al., 2011). This approach reveals that people rarely engage in exhaustive planning, instead employing approximate and bounded algorithms that selectively expand promising branches while pruning costly ones (Huys et al., 2012, 2015). Planning depth and precision adapt flexibly to uncertainty (Fan et al., 2024), working memory capacity (Ying et al., 2024), and effort costs (Callaway et al., 2022), consistent with a resource-rational trade-off between expected reward and cognitive expenditure (Lieder and Griffiths, 2020).

While choice data yields numerous insights, additional behavioral measures are often needed for fine-grained algorithmic distinctions. This is because different planning models can produce the same choice—often the best available option—despite relying on different computations. To mitigate this indeterminacy, planning models can also be compared by their ability to predict human reaction time, since it provides an additional constraint on the underlying computation. Reaction time thus serves as a key behavioral measure for distinguishing among planning algorithms (Callaway et al., 2024; Schwöbel et al., 2024). Solway and Botvinick (2015) demonstrated this approach by modeling planning as an evidence integration process and found that planning is best explained as parallel evidence accumulation over time with competition across independent paths.

Modeling the temporal dynamics of multi-stage decision making, however, poses a methodological challenge. Unlike sequential sampling models (SSMs) used in simple choices (Ratcliff, 1978; Brown and Heathcote, 2008; Turner et al., 2018), computational models of multi-stage decision making generally lack analytical solutions for the joint probability of choices and reaction times. As a result, researchers have fit models based on summary statistics, or used likelihood-free inference approaches, both of which have limitations. Fitting summary statistics compresses behavior into a few values (Solway and Botvinick, 2015), potentially hiding variability in the data. Approximating the likelihood using marginal RT quantiles (Heathcote et al., 2002) assumes independence across stages, an assumption often violated in multi-stage tasks. Inverse binomial sampling (IBS) provides unbiased likelihood estimates for discrete data (van Opheusden et al., 2020), but extending it to continuous variables is nontrivial, leaving detailed aspects of the data distribution unmodeled.

Here we develop a probability density approximation (PDA) method, building on Turner and Sederberg (2014), for estimating likelihoods in multi-stage decision making. We validate PDA using evidence accumulation models from Solway and Botvinick (2015) and a new human planning experiment, demonstrating that PDA closely matches analytical solutions where available and improve parameter recovery over alternative methods. When applied to human data, fitting the full joint distribution of reaction times—rather than summary statistics or marginal fits—reveals systematic model misspecification undetectable from psychometric curves alone and leads to different model selection conclusions. These results demonstrate that modeling the full reaction time distribution is critical for recovering planning dynamics and for reliably evaluating multi-stage decision models.

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

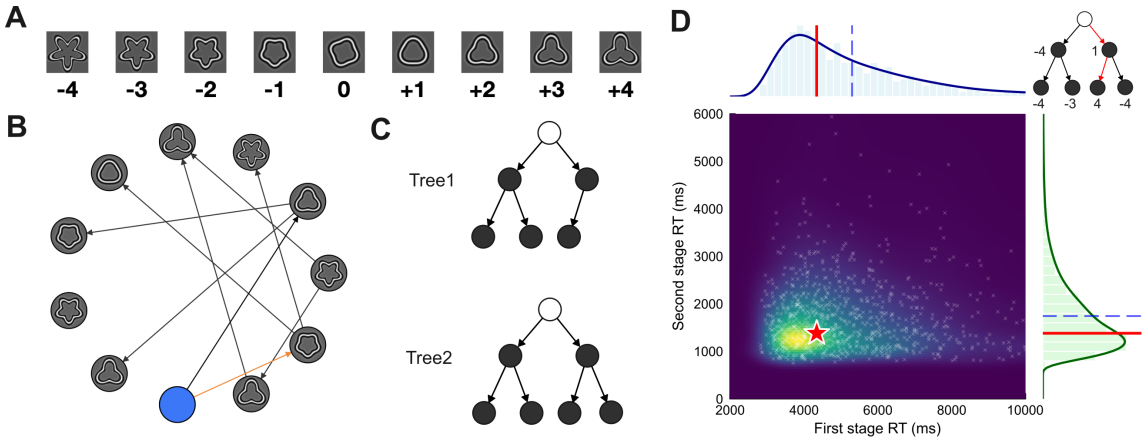

Figure 1: Task design and illustration of probability density approximation (PDA) method. (A) The reward associated with each shape ranged from -4 to 4. (B) The task interface shown to the participants. Participants select a node by key press. Their current state is highlighted in blue, and the chosen action is highlighted in orange. The goal is to maximize number of points by navigating in the graph. (C) Two tree configurations. Each node is one state, and each arrow corresponds to an action in the graph. (D) An illustration of PDA. The method draws samples from the simulator model and then uses a two-dimensional Gaussian kernel to estimate the joint probability density of response time (RT). The red lines and star indicate the true response time a participant had in this example trial. The blue dashed line indicates the mean of the response time distribution.

Methods

Decision tree navigation task

We adapted the sequential decision task from Solway and Botvinick (2015), modifying the spatial layout to prevent participants from using physical proximity as a cue to connectivity. States were arranged in a circle (Callaway et al., 2024; Correa et al., 2023; Zhu et al., 2022), which limits rapid visual scanning for high-reward regions and requires participants to internally track action sequences, thereby increasing reliance on internal planning. The task environment consisted of 11 states, each labeled with a shape indicating the points gained or lost upon visiting (Fig. 1A). Arrows indicated available transitions between states. On each trial, participants navigated from an initial state to a terminal state, accumulating points along their trajectory (Fig. 1B). To select an action, participants cycled through available options using the 'F' key and confirmed with the 'J' key; decisions were irreversible. Participants' goal was to maximize total points. Trials used one of two tree configurations—'Tree 1' or 'Tree 2'—(Fig. 1C), presented in random order.

Experimental procedure The experiment began with an interactive instruction phase introducing the task structure and rules. Participants then completed guided practice trials, including moving to designated locations to demonstrate task comprehension. To verify learning of the shape–reward associations, participants completed 10 binary-choice practice trials and were required to achieve the maximum points on all trials. Participants who did not meet this criterion received additional clarification before proceeding. Following practice, participants completed 200 or 300 experimental trials.

Participants Forty-five participants were recruited from Prolific who is fluent in English. For model fitting and analysis, we included only trials with the 'Tree 2' structure, which provides the minimum branching complexity needed to distinguish between the candidate models. Following Solway and Botvinick (2015), we excluded trials with response times shorter than 500 ms or longer than 10 s. After exclusions, the final dataset comprised 4589 trials.

Computational models

Solway and Botvinick (2015) proposed a family of 14 evidence accumulation models for model-based tree search, differing in their assumptions about noise structure, pruning mechanisms, and search strategies. For brevity, we introduce only the two models used to validate our likelihood-free inference method; see full model specifications in Solway and Botvinick (2015).

Forward greedy model (FG) The forward greedy (FG) model assumes that decisions at each stage are made independently, based only on immediate rewards. At the first stage, evidence for each action accumulates according to:

$$\begin{aligned} E_L^{t+1} &= E_L^t + d_1 \cdot R_L + \epsilon_L, \\ E_R^{t+1} &= E_R^t + d_1 \cdot R_R + \epsilon_R, \end{aligned} \tag{1}$$

where E denotes cumulative evidence for a given action, d_1 is the drift rate, R_L and R_R are rewards at the respective next states, and ϵ denotes zero-mean Gaussian noise $\epsilon \sim \mathcal{N}(0, \sigma^2 I)$ with standard deviation $\sigma = 0.01$. Following Solway and Botvinick (2015), we use L and R to denote the two actions, though these do not correspond to spatial locations given our circular layout. A decision occurs when the evidence difference exceeds threshold θ_1.

At the second stage, the two states within the chosen branch compete via an analogous process. For example, if action R is selected:

$$\begin{aligned} E_{RL}^{t+1} &= E_{RL}^{t} + d_2 \cdot R_{RL} + \epsilon_{RL}, \\ E_{RR}^{t+1} &= E_{RR}^{t} + d_2 \cdot R_{RR} + \epsilon_{RR}, \end{aligned} \tag{2}$$

with threshold θ_2. Each stage includes a non-decision time parameter (T_1 and T_2) accounting for perceptual and motor processes. Integrators reset to zero at each stage.

Two stage independent-path model (TS) In contrast to the stage-wise FG model, the two-stage independent-path (TS) model treats each complete path through the tree as a single evidence integrator, allowing information about future rewards to influence the first-stage decision. Evidence for each of the four paths is updated according to the cumulative reward along that path:

$$\begin{aligned} E_{L,L}^{t+1} &= E_{L,L}^{t} + (d_1 R_L + \epsilon_L) + (d_1 R_{LL} + \epsilon_{LL}), \\ E_{L,R}^{t+1} &= E_{L,R}^{t} + (d_1 R_L + \epsilon_L) + (d_1 R_{LR} + \epsilon_{LR}), \\ E_{R,L}^{t+1} &= E_{R,L}^{t} + (d_1 R_R + \epsilon_R) + (d_1 R_{RL} + \epsilon_{RL}), \\ E_{R,R}^{t+1} &= E_{R,R}^{t} + (d_1 R_R + \epsilon_R) + (d_1 R_{RR} + \epsilon_{RR}). \end{aligned} \tag{3}$$

A first-stage decision occurs when the difference between the best and second-best paths exceeds threshold θ_1. The second stage continues accumulating evidence for the two remaining paths until their difference exceeds θ_2, with $\theta_2 > \theta_1$ to ensure the second-stage decision follows the first. Non-decision time parameters T_1 and T_2 apply as in the FG model.

Model fitting

We compare three approaches for fitting models to behavioral data: probability density approximation (PDA), which we develop here for multi-stage decisions; residual sum of squares (RSS) (Solway and Botvinick, 2015); and inverse binomial sampling (IBS) (van Opheusden et al., 2020).

Probability density approximation (PDA) Our approach extends the PDA framework of Turner and Sederberg (2014) to multi-stage decision-making. The key idea is to approximate the likelihood of observed choices and reaction times by simulating the model many times and using kernel density estimation on the resulting samples. PDA requires two components: (1) a dataset $\mathcal{D} = \{\boldsymbol{s}_i, \boldsymbol{c}_i, \boldsymbol{t}_i\}_{i=1}^{N}$ consisting of N trials, each characterized by a state vector $\boldsymbol{s}_i$, choice vector $\boldsymbol{c}_i$, and reaction time vector $\boldsymbol{t}_i$; and (2) a generative model $\mathcal{M}$ that takes a state $\boldsymbol{s}$ and parameter vector $\boldsymbol{\theta}$ as input and outputs responses $\boldsymbol{r}$ over choices and reaction times.

Assuming independence across trials, the joint likelihood of choices and reaction times for a particular model is:

$$\mathcal{L}(\boldsymbol{\theta} \mid \mathcal{D}) = \prod_{i=1}^{N} \mathcal{M}(\boldsymbol{c}_i, \boldsymbol{t}_i \mid \boldsymbol{\theta}, \boldsymbol{s}_i) \tag{4}$$

Following Turner and Sederberg (2014), we factorize the joint distribution using the chain rule for trial i:

$$p(c_1, c_2, t_1, t_2 \mid \boldsymbol{\theta}, \boldsymbol{s}) = p(c_1, c_2 \mid \boldsymbol{\theta}, \boldsymbol{s})\; p(t_1, t_2 \mid c_1, c_2, \boldsymbol{\theta}, \boldsymbol{s}) \tag{5}$$

We use this identity to compute the empirical likelihood of choices and reaction times in our experimental dataset for each parameter/model combination. The choice probability, $p(c_1, c_2 \mid \boldsymbol{\theta}, \boldsymbol{s})$, is estimated from empirical frequencies:

$$p(c_1, c_2 \mid \boldsymbol{\theta}, \boldsymbol{s}) = \frac{n(c_1, c_2)}{J} \tag{6}$$

where $n(c_1, c_2)$ is the number of simulated samples yielding choice pair (c_1, c_2), and J is the total number of samples. The conditional RT density, $p(t_1, t_2 \mid c_1, c_2, \boldsymbol{\theta}, \boldsymbol{s})$, is estimated via multivariate kernel density estimation. Let $\boldsymbol{x} = (t_1, t_2)$ and let $\{\boldsymbol{x}_i\}_{i=1}^{n}$ be the simulated RTs matching the observed choice. The density estimate is:

$$\hat{f}_H(\boldsymbol{x}) = \frac{1}{n} \sum_{i=1}^{n} K_H(\boldsymbol{x} - \boldsymbol{x}_i), \tag{7}$$

where K_H is a Gaussian kernel with bandwidth matrix H:

$$K_H(\boldsymbol{u}) = \frac{1}{(2\pi)^{d/2} |H|^{1/2}} \exp\!\left(-\tfrac{1}{2} \boldsymbol{u}^\top H^{-1} \boldsymbol{u}\right) \tag{8}$$

The bandwidth matrix H uses a rule-of-thumb scaling based on the sample covariance matrix Σ, which balances bias and variance in the density approximation (Silverman, 1986):

$$H = \left(\frac{4}{d+2}\right)^{1/(d+4)} n^{-1/(d+4)} \Sigma \tag{9}$$

where d is the RT vector dimensionality and Σ is the sample covariance. As $n \to \infty$, $H \to 0$ and the estimate converges to the true density.

Figure 1D illustrates PDA for a single trial: the model is simulated repeatedly for a given reward configuration, and the resulting RT samples are used to estimate the joint density (Eq. 7). The observed RT (red marker) falls near the density peak, indicating high likelihood under the current parameters.

Residual sum of squares (RSS) In Solway and Botvinick (2015), for each candidate parameter vector $\boldsymbol{\theta}$, the model was simulated once for every empirical trial, and predictions were summarized into psychometric curves (i.e., choice accuracy and mean RT as functions of task difficulty). The objective function was the residual sum of squares (RSS) between empirical and simulated curves, after rescaling RTs to balance their contribution:

$$\mathrm{RSS}(\boldsymbol{\theta}) \;=\; \sum_{j=1}^{n} \left(y_j^{\mathrm{data}} - y_j^{\mathrm{sim}}(\boldsymbol{\theta})\right)^2, \tag{10}$$

where y_j^{data} and $y_j^{\mathrm{sim}}(\boldsymbol{\theta})$ are the j-th binned statistics from data and simulation, respectively.

Inverse binomial sampling (IBS) IBS estimates log-likelihoods by repeatedly simulating from the model until a simulated response matches the observed one. For continuous variables like RT, a match is defined when the difference between a simulated response $\tilde{r}_i$ and the observed response r_i falls within a threshold ε (van Opheusden et al., 2020):

$$\mathcal{L}_\varepsilon(\boldsymbol{\theta}) = \sum_{i=1}^{N} \log \Pr(|\tilde{r}_i - r_i| \le \varepsilon \mid \boldsymbol{\theta}, \boldsymbol{s}_i) \tag{11}$$

As $\varepsilon \to 0$, the estimate approaches the true likelihood under regularity conditions, but the expected number of required samples diverges, leading to a bias–variance tradeoff.

Optimization All three fitting methods require stochastic simulation, making objective function evaluation noisy and computationally expensive. We used Bayesian Adaptive Direct Search (BADS) (Acerbi and Ma, 2017; Singh and Acerbi, 2024), a derivative-free optimizer designed for expensive, potentially non-smooth objectives.

Validation of the PDA approach

Comparison with analytical solution To evaluate the accuracy of PDA under finite sampling, we compared it against an analytical likelihood available for the FG model. Because the FG model treats stages independently, each stage can be formulated as a drift–diffusion process:

$$dx = \mu dt + \sigma dW \tag{12}$$

where W is the standard Wiener process and the drift rate $\mu = d_i(R_L - R_R)$ is determined by the reward difference at each stage. This formulation yields closed-form first-passage time (FPT) densities for each stage (Drugowitsch, 2016):

$$\begin{aligned} g_1 &= f_{\text{FPT}}(t_1 \mid \mu_1, \theta_1, \text{upper}_1), \\ g_2 &= f_{\text{FPT}}(t_2 \mid \mu_2, \theta_2, \text{upper}_2), \end{aligned} \tag{13}$$

with total log-likelihood:

$$\log p(t_1, t_2 \mid \boldsymbol{c}, \boldsymbol{\theta}) = \log g_1 + \log g_2. \tag{14}$$

To compare analytical and PDA likelihoods, we first fitted the FG model using the analytical solution, then evaluated trial-wise PDA log-likelihoods at the fitted parameters using $J = 1000$ simulations, requiring at least 100 samples with matching choices. We found that PDA closely matches analytical log-likelihoods, exhibiting high correlation and minimal deviations across the likelihood range (Fig. 2A-B). PDA also captures the local geometry relevant for optimization when we varied θ_1 locally around the optimum and fixed all parameters at their fitted values (Fig. 2C). Finally, subject-level root mean square error (RMSE) are uniformly low, indicating stable agreement across participants (Fig. 2D).

Parameter recovery Parameter recovery assesses whether an inference method can reliably recover true generative parameters under controlled conditions (Wilson and Collins, 2019). We used this approach to evaluate PDA relative to RSS and IBS, with analytical likelihood as a benchmark.

We generated 50 synthetic datasets of 100 trials each from the FG model. For each dataset, parameters were sampled from the following distributions: d_1 and d_2 log-uniformly from $[10^{-5}, 10^{-3}]$, θ_1 and θ_2 uniformly from $[0.01, 1]$, and T_1 and T_2 uniformly from $[100, 5000]$ ms. All six parameters were then jointly refit using each method.

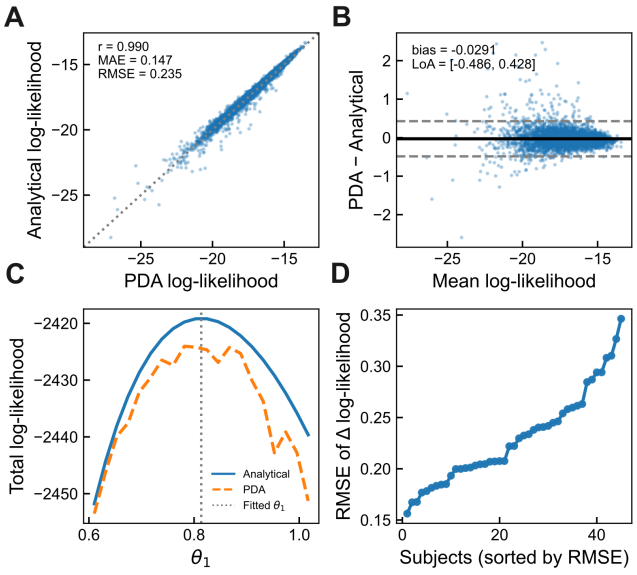

Figure 2: Comparison between analytical solution and PDA likelihood estimates in FG model. (A) Trial-wise PDA and analytical log-likelihood show strong correlation. (B) Difference indicates near-zero bias and tight agreement, with no systematic deviation across the likelihood range. (C) A one-parameter likelihood slice around the fitted optimum shows that PDA preserves the local shape and peak of the likelihood landscape. (D) Subject-level RMSE of trial-wise log-likelihood differences remains low across participants.

Recovery quality differed markedly across methods (Fig. 3). PDA achieved near-perfect recovery comparable to the analytical benchmark, with fitted parameters tightly aligned to the identity line. RSS showed substantial variability and systematic deviations, reflecting poor identifiability. IBS performed comparably to PDA in correlation but systematically underestimated decision thresholds, likely because its tolerance-based matching criterion does not fully capture the right-skewed shape of RT distributions.

Results

Inconsistent model selection across methods

We first aimed to replicate the key finding from Solway and Botvinick (2015)—that the TS model best explains human planning behavior. Using the original RSS fitting procedure, we simulated all 14 candidate models from that study on our new experimental data, aggregating predictions into group-level psychometric curves and minimizing RSS between empirical and simulated curves. Consistent with the original findings, the TS model provided the best fit to both accuracy and reaction time psychometric curves [1] at the first stage and second stage (Fig. 4). Model comparison using their Bayesian information criterion (BIC) favored the TS model over alternatives (TS: BIC = -289.11; FG: BIC = -257.96).

[1] For brevity, we only plot TS and FG models via RSS and PDA.

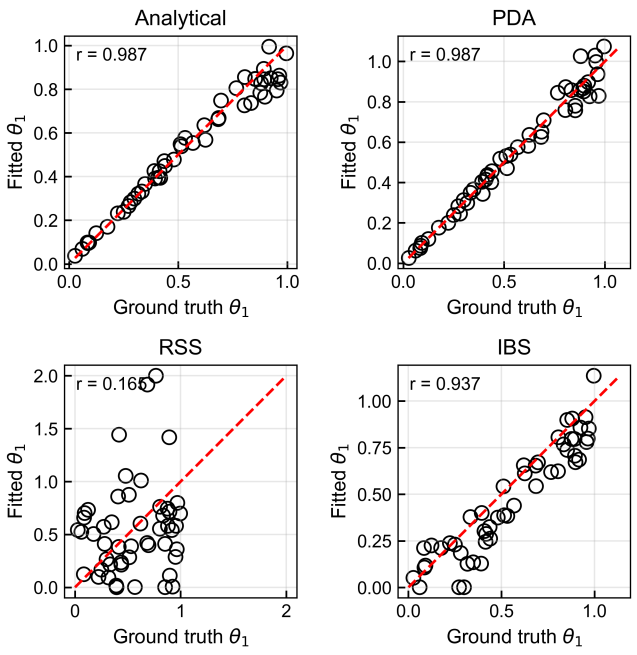

Figure 3: Parameter recovery across different fitting method. Analytical likelihood, PDA, RSS and IBS are shown as scatter plots comparing ground truth θ_1 with fitted θ_1 in forward greedy model. Dashed lines indicate the identity line, and r indicates Pearson correlation coefficient.

We next applied our PDA method by fitting model parameters at the individual level to assess the robustness of the results. In contrast to the previous RSS result, the forward greedy model emerged as the preferred model under individual-level fitting (TS: BIC = 3556.85; FG: BIC = 3523.47). As a reference, IBS also prefers FG over TS (TS: BIC = 446.68; FG: BIC = 432.79). To further examine this discrepancy, we simulated data using individually fitted parameters and constructed group-level psychometric curves. Although both models underestimate choice accuracy at low difficulty levels in both stages relative to human data (Fig. 4A and Fig. 4B), the most pronounced divergence arises in the second-stage reaction times as a function of second-stage reward difference (Fig. 4E), where the TS model differs substantially in its fit.

Model misspecification from posterior check

To elucidate the source of the divergent model comparison results and the qualitative differences observed in the psychometric curves, we draw on insights from Ritz et al. (2026), who demonstrated that apparent model mimicry can arise artefactually from model misspecification. Accordingly, we conducted posterior predictive checks using parameters fitted with RSS and PDA, examining the reaction time distributions at both the first and second stages via kernel density estimation. As shown in Fig. 5, although the TS model fitted using RSS best reproduces the aggregate psychometric curves, it systematically deviates from the empirical reaction time distributions. This discrepancy is particularly pronounced at the second stage, where the model predicts an unrealistically sharp peak rather than the right-skewed distribution observed in human data. In contrast, while PDA does not fully capture all fine-grained features of the second-stage reaction time pattern (in Fig. 4E), it more accurately reproduces the overall distribution (Fig. 5). For reference, the reaction time distribution at the second stage is also not well accounted for by the FG model using either RSS or PDA.

Why does the bad prediction at the second stage reaction time distribution only occur in RSS instead of PDA? A plausible explanation related to the assumption in the TS model that evidence continues accumulating between stages. In this model, the first-stage decision threshold is determined solely by the difference between the best and second-best paths. Consequently, upon entering the second stage, the evidence difference within the selected branch may already exceed the second-stage threshold, resulting in near-immediate termination[2] and a small reaction time. This happen in 57.52% of trials when using RSS, but only 24.73% of trials when using PDA. Rapid terminations occur more frequently when second-stage reward differences are large and when the best and second-best paths are in different branches. Because RSS optimizes only mean reaction times and ignores distributional variance, it disproportionately exploits these rapid terminations to account for decreases in second-stage mean reaction time with reward difference. In contrast, PDA penalizes assigning negligible probability mass to regions away from the empirical distribution, preventing such degenerate solutions from maximizing the likelihood and thereby yielding more realistic reaction time distributions.

The rapid termination mechanism in the TS model also results in inferior fits relative to the FG model under likelihood-based estimation methods such as PDA and IBS. Unlike RSS, which primarily matches mean statistics, PDA accounts for higher-order properties of the reaction time distribution, including variance and skewness, making it insufficient to concentrate probability mass solely around the mean for each difficulty level. To accommodate the broader variability observed in second-stage reaction times, the TS model compensates by substantially reducing its second-stage drift rate d_2, yielding estimates approximately an order of magnitude smaller than those of the FG model under PDA (TS: $d_2 = 7.53 \times 10^{-5}$; FG: $d_2 = 7.81 \times 10^{-4}$). This adjustment produces excessively long and highly skewed predicted reaction time distributions, leading to systematic overestimation of second-stage reaction times (Fig. 4E). In contrast, the FG model affords greater flexibility in capturing second-stage dynamics, as it is not constrained by first-stage, and therefore provides a comparatively better fit under likelihood-based methods, despite remaining limitations in fully capturing the posterior distribution.

[2]Near-immediate termination is defined as a second-stage decision occurring within 50 ms except non-decision time.

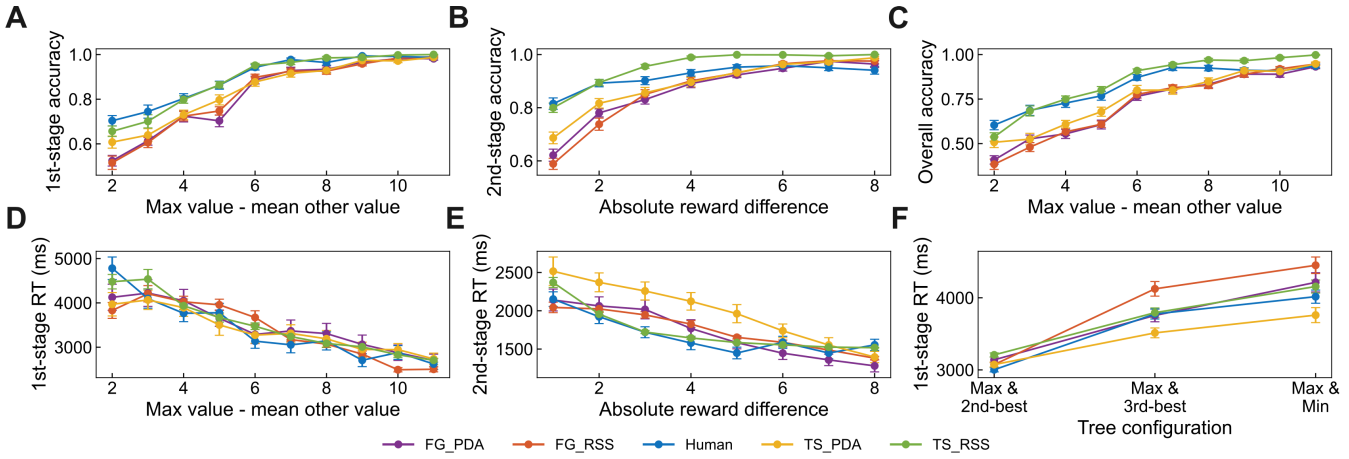

Figure 4: Replication of psychometric curves in Solway and Botvinick (2015). Fits of the TS and FG model using RSS and PDA are shown. (A) First stage choice accuracy as a function of the difference between the maximum path value and the average of other path values. (B) Second stage choice accuracy as a function of the absolute reward difference. (C) Overall accuracy as a function of the difference between the maximum path value and the average of other path values. (D) First stage reaction time versus the difference between the maximum path value and the average of other path values. Only correct trials are included. (E) Second stage reaction time versus absolute reward difference. (F) First stage reaction time across different tree configuration.

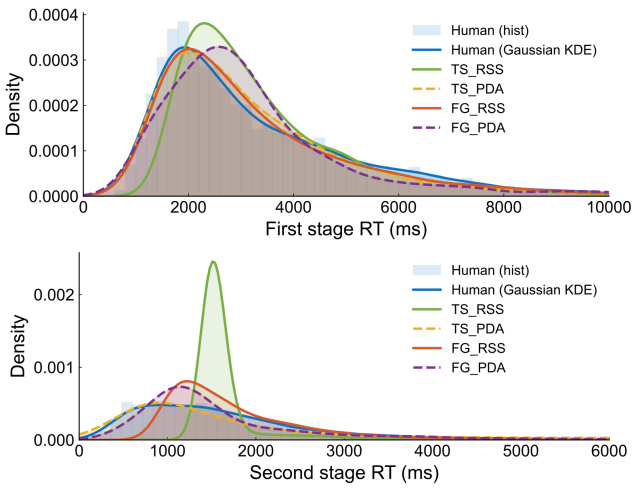

Figure 5: Posterior predictive checks of first-stage and second-stage reaction time distributions for the TS and FG models fitted using RSS and PDA.

Discussion

In this paper, we developed a probability density approximation (PDA) method for fitting cognitive models with continuous variables such as reaction time in multi-stage decision making. Validation against analytical likelihoods and parameter recovery analyses demonstrated that PDA provides robust and reliable estimates, outperforming both summary-statistic (RSS) and tolerance-based (IBS) approaches. Applied to human data, PDA revealed that in our task, model selection conclusion about planning, such as whether it proceeds via parallel integration across independent paths (Solway and Botvinick, 2015), depends critically on the fitting method: when full RT distributions are modeled rather than summary statistics, a simpler forward greedy model is preferred.

This reversal underscores a broader methodological lesson: good agreement with psychometric curves does not guarantee that a model adequately captures human behavior. In our task, fitting only mean behavior while neglecting distributional shape proved misleading, as model-specific assumptions (specifically, the rapid second-stage terminations in the TS model) artificially improved fits to mean RT and led to incorrect conclusions. PDA facilitates more principled model evaluation by incorporating full distributional information, revealing misspecification that summary statistics conceal. Moreover, individual-level fitting becomes feasible even when the number of trials per participant is limited, enabling examination of individual differences in planning strategies.

Several limitations warrant consideration. First, the accuracy of PDA likelihood estimates depends critically on the number of simulated samples used for kernel density estimation. As the number of decisions increase, the state space and the dimensionality of the joint reaction time distribution grow rapidly, making likelihood estimation more computationally demanding. Indeed, small sample sizes can introduce substantial bias, particularly for low-probability events. Future work could address sample efficiency by incorporating importance sampling (Tran et al., 2020) or using neural networks to approximate likelihood (Fengler et al., 2021). Second, prior work has shown that omission trials can substantially affect parameter estimates (Leng et al., 2024); integrating omissions into our framework is another promising direction. Last, neither model we tested fully captured human RT distributions, suggesting that more flexible architectures may be required.

Acknowledgments

The authors thank the anonymous reviewers and members of the Mattar Lab for their helpful comments and feedback on the project.

References

Acerbi, L. and Ma, W. J. (2017). Practical bayesian optimization for model fitting with bayesian adaptive direct search.

Brown, S. D. and Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, 57(3):153–178.

Callaway, F., Van Opheusden, B., Gul, S., Das, P., Krueger, P. M., Griffiths, T. L., and Lieder, F. (2022). Rational use of cognitive resources in human planning. *Nature Human Behaviour*, 6(8):1112–1125.

Callaway, F., Yu, M., and Mattar, M. G. (2024). Revealing human planning strategies with eye-tracking. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 46(0).

Correa, C. G., Ho, M. K., Callaway, F., Daw, N. D., and Griffiths, T. L. (2023). Humans decompose tasks by trading off utility and computational cost. *PLoS Computational Biology*, 19(6):e1011087.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., and Dolan, R. J. (2011). Model-Based Influences on Humans' Choices and Striatal Prediction Errors. *Neuron*, 69(6):1204–1215.

Drugowitsch, J. (2016). Fast and accurate monte carlo sampling of first-passage times from wiener diffusion models. *Scientific Reports*, 6(1):20490.

Fan, H., Callaway, F., and Gershman, S. J. (2024). Uncertainty-driven exploration during planning.

Fengler, A., Govindarajan, L. N., Chen, T., and Frank, M. J. (2021). Likelihood approximation networks (LANs) for fast inference of simulation models in cognitive neuroscience. *eLife*, 10:e65074.

Heathcote, A., Brown, S., and Mewhort, D. J. K. (2002). Quantile maximum likelihood estimation of response time distributions. *Psychonomic Bulletin & Review*, 9(2):394–401.

Huys, Q. J. M., Eshel, N., O'Nions, E., Sheridan, L., Dayan, P., and Roiser, J. P. (2012). Bonsai Trees in Your Head: How the Pavlovian System Sculpts Goal-Directed Choices by Pruning Decision Trees. *PLoS Computational Biology*, 8(3):e1002410.

Huys, Q. J. M., Lally, N., Faulkner, P., Eshel, N., Seifritz, E., Gershman, S. J., Dayan, P., and Roiser, J. P. (2015). Interplay of approximate planning strategies. *Proceedings of the National Academy of Sciences*, 112(10):3098–3103.

Kuperwajs, I., Russek, E. M., Mattar, M. G., Ma, W. J., and Griffiths, T. L. (2025). Looking deeper into the algorithms underlying human planning. *Trends in Cognitive Sciences*, 0(0).

Leng, X., Fengler, A., Shenhav, A., and Frank, M. J. (2024). The Perils of Omitting Omissions when Modeling Evidence Accumulation. *Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)*.

Lieder, F. and Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, 43:e1.

Mattar, M. G. and Lengyel, M. (2022). Planning in the brain. *Neuron*, 110(6):914–934.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85(2):59–108.

Ritz, H., Frömer, R., and Shenhav, A. (2026). Misspecified models create the appearance of adaptive control during value-based choice. *Communications Psychology*, 4(1):11.

Schwöbel, S., Marković, D., Smolka, M. N., and Kiebel, S. (2024). Joint modeling of choices and reaction times based on bayesian contextual behavioral control. *PLOS Computational Biology*, 20(7):e1012228.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Routledge, London: Chapman & Hall.

Singh, G. S. and Acerbi, L. (2024). PyBADS: Fast and robust black-box optimization in python. *Journal of Open Source Software*, 9(94):5694.

Solway, A. and Botvinick, M. M. (2015). Evidence integration in model-based tree search. *Proceedings of the National Academy of Sciences*, 112(37):11708–11713.

Tran, M.-N., Scharth, M., Gunawan, D., Kohn, R., Brown, S. D., and Hawkins, G. E. (2020). Robustly estimating the marginal likelihood for cognitive models via importance sampling. *Behavior Research Methods*, 53(3):1148–1165.

Turner, B. M., Schley, D. R., Muller, C., and Tsetsos, K. (2018). Competing theories of multialternative, multiattribute preferential choice. *Psychological Review*, 125(3):329–362.

Turner, B. M. and Sederberg, P. B. (2014). A generalized, likelihood-free method for posterior estimation. *Psychonomic Bulletin & Review*, 21(2):227–250.

van Opheusden, B., Acerbi, L., and Ma, W. J. (2020). Unbiased and efficient log-likelihood estimation with inverse binomial sampling. *PLOS Computational Biology*, 16(12):e1008483.

Wilson, R. C. and Collins, A. G. (2019). Ten simple rules for the computational modeling of behavioral data. *eLife*, 8:e49547.

Ying, Z., Callaway, F., Kiyonaga, A., and Mattar, M. G. (2024). Resource-rational encoding of reward information in planning. *Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)*.

Zhu, S., Lakshminarasimhan, K. J., Arfaei, N., and Angelaki, D. E. (2022). Eye movements reveal spatiotemporal dynamics of visually-informed planning in navigation. *Elife*, 11.