PLOS Computational Biology

RESEARCH ARTICLE

Decomposing response inhibition: A POMDP model

Wenting Wang[1,2,3*], **Tobias Kaufmann**[2,4,5], **Peter Dayan**[1,4,6]

1 Department of Computational Neuroscience, Max Planck Institute for Biological Cybernetics, Tübingen, Germany, **2** Department of Psychiatry and Psychotherapy, Tübingen Center for Mental Health, University of Tübingen, Tübingen, Germany, **3** International Max Planck Research School (IMPRS) for The Mechanisms of Mental Function and Dysfunction (MMFD), Graduate Training Centre of Neuroscience (GTC), Tübingen, Germany, **4** German Center for Mental Health (DZPG), Tübingen, Germany, **5** Centre for Precision Psychiatry, Institute of Clinical Medicine, University of Oslo, Oslo, Norway, **6** Department of Informatics, University of Tübingen, Tübingen, Germany

* wenting.wang@tuebingen.mpg.de

Check for updates

OPEN ACCESS

Citation: Wang W, Kaufmann T, Dayan P (2026) Decomposing response inhibition: A POMDP model. PLoS Comput Biol 22(9): e1014068. https://doi.org/10.1371/journal.pcbi.1014068

Editor: Jian Liu, University of Birmingham, UNITED KINGDOM OF GREAT BRITAIN AND NORTHERN IRELAND

Received: February 27, 2026

Accepted: September 10, 2026

Published: September 25, 2026

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.pcbi.1014068

Copyright: © 2026 Wang et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Inhibitory control is a core cognitive function whose competence varies across the population, with impairments often observed in psychiatric conditions such as attention deficit hyperactivity disorder (ADHD). The Stop Signal Task (SST) is a widely used paradigm for assessing this ability. However, conventional formalizations of SST performance, such as the independent race model, rely on assumptions that are frequently violated in modern experimental designs. Furthermore, they typically fit only mean reaction times, overlooking crucial trial-by-trial dynamics. To address these limitations, we formalize the SST as a partially observable Markov decision process (POMDP). This framework characterizes inhibitory control with two components: noisy perceptual inference regarding stimuli and optimal control balanced against potential costs. To fit this model to the Adolescent Brain Cognitive Development (ABCD) study baseline cohort (N = 3,567), we introduce Transformer-encoded Simulation-Based Inference (TeSBI). This end-to-end architecture learns compact, sequence-aware embeddings from raw behavioral data. It enables efficient, amortized inference of individual-level posteriors. Extensive validation confirms it extracts reliable and identifiable parameters. We identify distinct latent computational attributes associated with scores on ADHD questionnaires. Controlling for sex, IQ, and medication status, children with higher ADHD scores exhibit subtle but robust shifts in computational attributes. They show a reduction in go cue directional precision, alongside a blunted sensitivity to stop error, go error and time costs. The learned embedding space reveals a continuous manifold in which children with higher ADHD scores are heterogeneously distributed, rather than forming distinct disorder clusters. This indicates that similar clinical characteristics can emerge from diverse combinations of computational mechanisms, supporting a dimensional perspective on neurodiversity. Our end-to-end framework can be extended to a broader range of

Data availability statement: Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive Development (ABCD) Study (https://abcdstudy.org). To protect participant confidentiality and privacy, these clinical datasets are restricted. Access to the raw and processed data (Data Release 5.1) can be obtained by researchers through the NIMH Data Archive (NDA) (https://nda.nih.gov/abcd), and ABCD data can now also be found at the NBDC Data Hub (https://www.nbdc-data-hub.org/). A listing of participating sites and a complete listing of the study investigators can be found at https://abcdstudy.org/consortium_members/. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in the analysis or writing of this report. This study reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators. All custom code used for the POMDP modeling, parameter inference, and statistical analyses is publicly available on the Open Science Framework (OSF) at https://doi.org/10.17605/OSF.IO/UQNCE and on GitHub at https://github.com/wenting-wang/sst-pomdp. To ensure full transparency and facilitate reproducibility, we have provided mock datasets and summary files in the repositories. This allows users to test the code and run the analysis pipelines without needing access to the restricted clinical dataset.

Funding: This work was supported by the Else Kröner-Fresenius-Stiftung (Else Kröner Medical Scientist Kolleg ClinBrAIn) awarded to W.W., T.K., and P.D.; the Max-Planck-Gesellschaft awarded to P.D.; and the Alexander von Humboldt-Stiftung awarded to P.D. The ABCD Study data used in this report is supported by the National Institutes of Health and additional federal partners under award numbers U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134, U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123, U24DA041147. A full list of supporters is available at https://abcdstudy.org/federal-partners.html. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the study.

Competing interests: The authors have declared that no competing interests exist.

cognitive tasks. It offers a scalable, theory-driven solution for analyzing large-scale behavioral data.

Author summary

Inhibitory control is essential for adjusting thoughts and behavior and is often impaired in conditions like ADHD. Traditional assessments often oversimplify the underlying decision-making. We address this using a biologically grounded mathematical framework that separates how we perceive signals from how we execute control. This framework remains valid in diverse experimental designs where traditional models fail. To apply this complex model at scale, we develop a specialized machine learning approach (TeSBI) to reverse-engineer individual cognitive profiles. Analyzing data from over 3,000 children, we linked higher ADHD scores to subtle, consistent computational shifts: a reduction in go cue directional precision, a blunted sensitivity to the costs of making mistakes or taking too much time. These alterations remain robust regardless of sex, IQ, or medication status. Rather than forming a single disorder cluster, these children displayed diverse cognitive profiles, supporting a dimensional view of neurodiversity. By combining theory-driven cognitive modeling with scalable data-driven inference, our framework effectively captures complex inhibitory mechanisms and enables the precise analysis of large-scale behavioral datasets. This approach paves the way for more personalized strategies in computational psychiatry by recognizing the inherent heterogeneity within clinical characteristics.

Introduction

Inhibitory control is a fundamental cognitive function that regulates thoughts, actions, and emotional responses. It is essential for flexible and goal-directed behavior in dynamic environments. This capacity is often impaired in psychiatric conditions such as attention deficit hyperactivity disorder (ADHD), autism spectrum disorder, and obsessive-compulsive disorder [1–3].

One popular method for examining individual differences in inhibitory control is the Stop Signal Task (SST) [4,5]. In this task, participants must make a fast *go* response to a cue unless a subsequent stop signal countermands the action. The interval between the go cue and the stop signal is the Stop Signal Delay (SSD). A shorter SSD makes inhibition easier. The SSD is typically adjusted on a trial-by-trial basis using a tracking algorithm (a *staircase*) to maintain a fixed proportion of inhibition failures. Standard SSTs are commonly modeled using evidence accumulation frameworks, such as the race model [4–6] or the drift diffusion model [7,8]. The parameters derived from these models provide significant insights into various group differences [9–12].

However, these standard models typically rely on subject-level aggregate metrics (e.g., mean reaction time, accuracy), which fail to capture the full complexity of behavioral dynamics. Furthermore, they assume that go and stop signals are processed independently. Neural evidence contradicts this assumption even in the standard SST [13,14], and some modern experimental designs directly violate it. A prominent example is the SST in the Adolescent Brain Cognitive Development (ABCD) study [15], where the stop signal visually masks the preceding go cue. In such contexts, race models relying on an assumption that go and stop signals are independent lead to biased estimates of the Stop Signal Reaction Time (SSRT) [16,17].

To address these limitations, we propose a general computational framework grounded in the Partially Observable Markov Decision Process (POMDP) formalism [18–23]. This framework naturally accommodates interactive go and stop processes while unifying perceptual inference and optimal control. By explicitly parameterizing key mechanisms, such as sensory precision, intrinsic costs, and probabilistic action selection, the POMDP model captures a wide range of behavioral variability across a heterogeneous population.

Fitting this model to the ABCD dataset presents computational challenges due to intractable likelihood functions, the adaptive SSD staircase, and sequential effects across trials. To overcome these obstacles, we introduce Transformer-encoded Simulation-Based Inference (TeSBI) [24–26]. Unlike traditional methods that rely on handcrafted summary statistics, the Transformer encoder learns compact, sequence-aware representations of behavioral data, preserving temporal dependencies and contextual information. A normalizing flow [27] enables an efficient and reliable estimation of individual-level parameter posteriors. Together, this end-to-end framework facilitates scalable model fitting for large-scale datasets.

In this study, we apply the POMDP-TeSBI framework to the ABCD SST dataset. Following rigorous model selection and comprehensive validation, we characterize individual perceptual capabilities and intrinsic valuation. We report the resulting subject-level parameter distributions and their small, but significant, associations with scores on ADHD questionnaires. Finally, we visualize distinct cognitive profiles using the embedding space learned by the Transformer encoder.

Materials and methods

Ethics statement

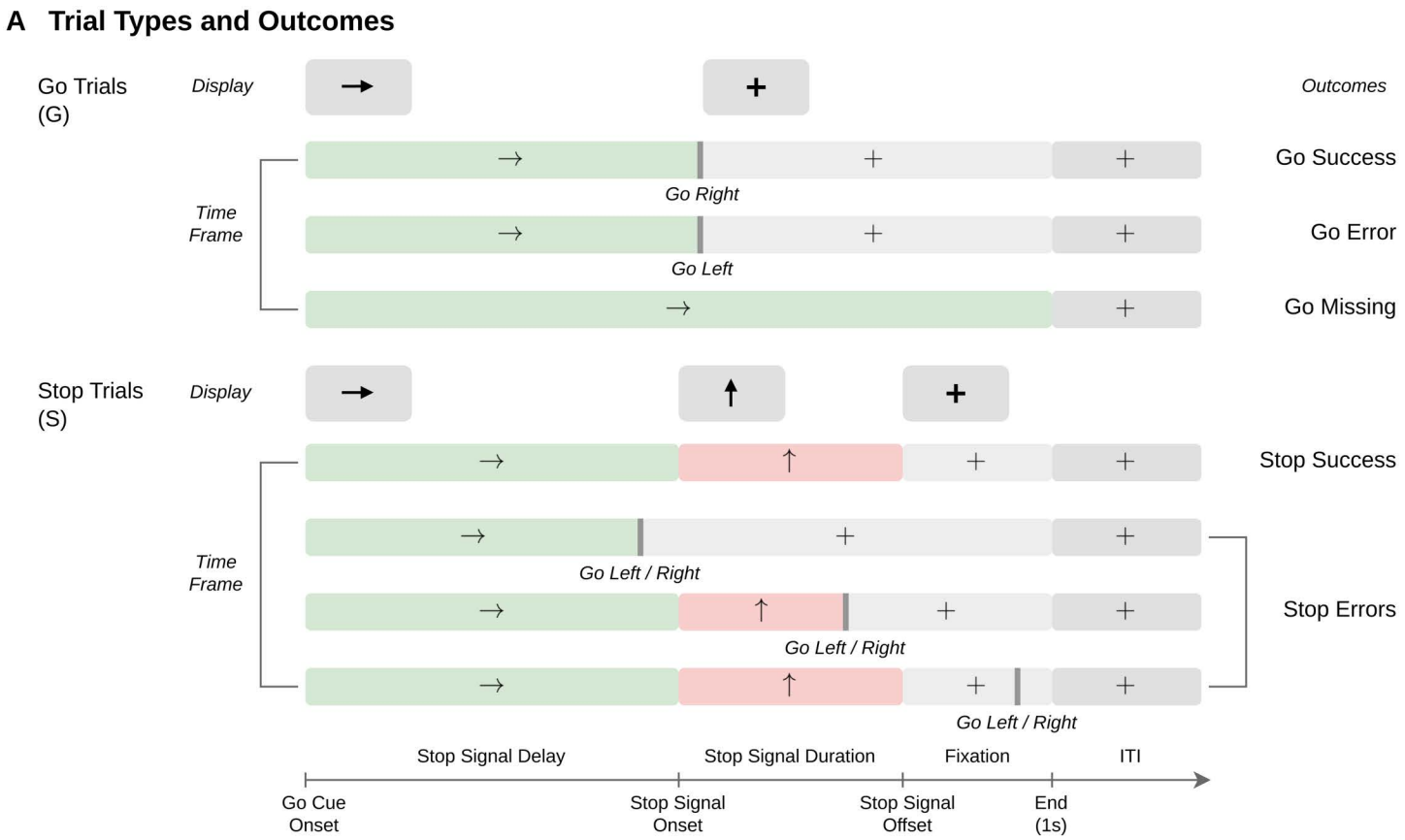
Study protocols have been approved by either local institutional review boards (IRB) or by reliance agreements with the central IRB at the University of California [28]. Written informed consent was obtained from the parents or legal guardians of all child participants, and written assent was obtained from the children themselves, prior to participation [29].

Experimental task and mental health measures

Stop signal task in the ABCD study. To examine inhibitory control, we analyze data from the Stop Signal Task (SST) deployed in the Adolescent Brain Cognitive Development (ABCD) Study, Annual Release 5.1 [15].

SST experimental designs are generally categorized as independent or dependent. In an independent design, the go cue remains visible when the stop signal appears. Classic formalizations, such as the independent race model [5], duly assume that go and stop signals are processed independently (although some neural evidence weighs against this assumption [13,14]). In contrast, the ABCD study utilizes a dependent design (Fig 1). In this version, the stop signal visually replaces, or masks, the preceding go cue. This replacement would disrupt and slow the processing of the go signal specifically on stop trials. Consequently, the core race model assumption that the go process remains identical across trial types is violated. This structural dependence confounds standard SSRT calculations [17].

Participants were instructed to press a button corresponding to the direction of a go cue (a left or right arrow presented with equal probability) as quickly and accurately as possible. They were instructed to withhold this response if a stop signal (an upward-pointing arrow) subsequently appeared. The session consisted of 360 trials divided into two blocks of 180 trials each. One-sixth of these (60 trials) were stop trials.



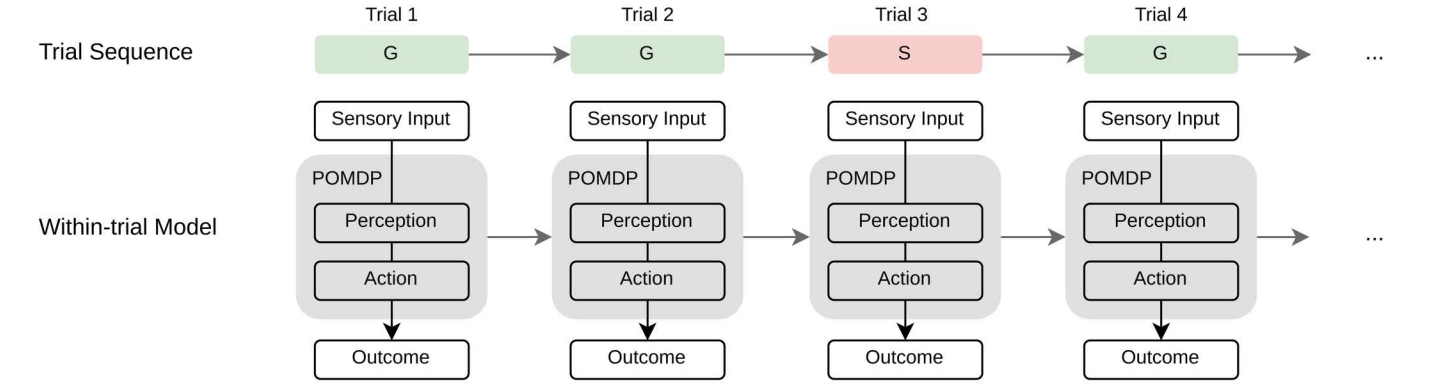


Fig 1. ABCD stop signal task and POMDP framework. (A) The task consists of 360 trials (300 go trials, 60 stop trials). Each trial has a fixed duration of 1000 ms. Black stimuli appear on a mid-gray background (top rectangles). Colored bars below illustrate the trial timing. When a response is registered, a fixation cross replaces the active stimulus for the remainder of the trial. On stop trials, a response registered at any point is categorized as a stop error. This includes responses during the initial Stop Signal Delay (SSD), the stop signal duration, or the subsequent fixation phase. The SSD is adjusted on a trial-by-trial basis to maintain a stop success rate of approximately 50%. **(B)** The POMDP framework models the within-trial dynamics of perception and choice. It is fit to the full behavioral dynamics, including the SSD staircase. The trial sequence shown is illustrative. Actual ABCD sessions follow predefined pseudo-randomized orders.

https://doi.org/10.1371/journal.pcbi.1014068.g001

All trials had a fixed duration of 1000 ms. They were separated by an inter-trial interval (ITI) jittered between 700 and 2000 ms. On go trials, the go cue remained on the screen until a response occurred or until a 1000-ms timeout was reached. Following a response, a fixation cross appeared for the remainder of the trial. On stop trials, the go cue appeared for a variable Stop Signal Delay (SSD). The stop signal then replaced the go cue for 300 ms, or until the 1000 ms trial limit if the SSD was $\geq$ 700 ms. If the participant responded despite the stop signal (i.e., failed inhibition), the stimulus was immediately replaced by a fixation cross.

The SSD was adjusted on a trial-by-trial basis using a tracking algorithm. This algorithm was designed to maintain a stop success rate of approximately 50%. The initial SSD was set to 50 ms. It increased by 50 ms following a successful inhibition and decreased by 50 ms following a failed inhibition. The SSD was intended to be constrained within a range of 50–850 ms.

We focused on baseline data from the ABCD study sample. We applied rigorous quality control criteria that evaluated both the technical experimental setup and participant task compliance. These inclusion criteria required: (1) at least three events for all trial types and outcomes; (2) no box switching due to reversed hardware mapping; (3) no task coding errors disrupting the SSD tracking algorithm; (4) fewer than four stop trials with a 0 ms SSD (these were typically caused by technical difficulties); (5) a mean reaction time (RT) on stop error trials not exceeding the mean RT on correct go trials; (6) a complete session of 360 trials, including exactly 60 stop trials; (7) fewer than four extremely fast trials (RT $\leq$ 200 ms); (8) a go missing rate $\leq$ 10%; (9) a go error rate $\leq$ 10%; and (10) no missing or incomplete demographic data.

After applying these exclusions, the final sample consisted of 3,567 participants. Further details regarding data quality control and model-agnostic statistical analyses are provided in Section F of S1 Appendix.

Mental health measures. At the baseline year (children aged 9–10 years), ADHD symptomatology was assessed using the parent-reported Child Behavior Checklist (CBCL). We utilized the DSM-5 ADHD raw scores [30]. These scores range from 0 to 14. Higher values indicate greater symptom severity. A continuous measure of ADHD symptomatology provides a more nuanced characterization of individual differences than binary diagnostic categories. This dimensional approach aligns with the National Institute of Mental Health's Research Domain Criteria (RDoC) framework [31]. Further details regarding CBCL administration are available in the ABCD Study protocols.

Model framework: POMDP

Consistent with established frameworks for perceptual decision-making [18,21,22,32], we formulate the SST as a Partially Observable Markov Decision Process (POMDP). We model it using two coupled observational streams: one for the directional go cue and one for the inhibitory stop signal. These streams are integrated via Bayesian inference to generate evolving belief states regarding the trial direction (Left versus Right) and trial type (Go versus Stop).

At each time step, based on these beliefs, the agent selects an action (Go-Left, Go-Right, or Wait) derived from a policy that optimizes total return. By balancing the costs of temporal delay, directional errors, missed responses, and failed inhibition, the model unifies Bayesian inference with optimal control. All model parameters are defined in Table 1 and detailed below.

Our POMDP framework accommodates different SST experimental designs. The main text focuses on a *dependent* design such as the ABCD task. As noted, in this design, the stop signal visually masks the go cue, causing their sensory observation processes to interact. The model also generalizes to *independent* designs. In these designs, the go and stop stimuli remain simultaneously visible without perceptual interference. Specifications for modeling independent SSTs are detailed in Section A.2 of S1 Appendix.

Perceptual inference. We follow standard Bayesian treatments of perceptual processing [22,32]. On each trial ν, a generative process specifies how latent states produce observable sensory inputs at each time step t. The inference process then uses these observations o to update beliefs about the latent states (Fig 2). We discretize the 1-second trial duration into $T = 40$ time steps of 25 ms each. For brevity, we omit the trial superscript ν where contextually clear.

Table 1. POMDP model parameters and descriptions.

Parameter	Name	Description
Sensory Parameters		
χ'	Go null	Probability of receiving uninformative noise regarding the go cue (sensory lapse/ambiguity).
χ	Go precision	Probability of successfully resolving the go cue direction (sensory fidelity).
δ'	Stop null	Probability of receiving uninformative noise regarding the stop signal (sensory lapse/ambiguity).
δ	Stop precision	Probability of successfully resolving the presence or absence of the stop signal (signal fidelity).
Cost Parameters		
c_{ge}	Go error cost	Subjective penalty for committing a directional error on go trials.
c_{gm}	Go missing cost	Subjective penalty for failing to respond (missing) on go trials.
c_{se}	Stop error cost	Subjective penalty for failing to inhibit the response on stop trials.
c_t	Time cost	Subjective penalty for each time step spent waiting.
Control and Dynamic Parameters		
φ	Inverse temperature	Degree of determinism in softmax policy.
λ	Hazard rate	Temporal prior of stop signal onset.
τ	Non-decision time	Sensory transmission and motor execution delays.
$\tilde{r}^{\nu}$	Stop prior	Predicted prior probability that the trial ν will be a stop trial.

https://doi.org/10.1371/journal.pcbi.1014068.t001

The stop process: generative and inference models In the generative model, the latent state $s^{\nu} \in \{0, 1\}$ indicates the true trial type (0 for Go trial, 1 for Stop trial). On stop trials, the stop signal appears at a specific stop signal delay (SSD, θ^{ν}) and lasts for $\mu + 1 = 12$ time steps (300 ms). The true stop signal state at time t is denoted as $z_t^{\nu} \in \{0, 1\}$ (0 for Absence, 1 for Presence). Participants receive a sensory observation $y_t^{\nu} \in \{A, P, N\}$ (Absence, Presence, or Null), generated from a multinomial distribution conditioned on z_t^{ν}. As detailed in Table 2, these observation likelihoods shift across the pre-signal, signal-on, and post-signal phases.

The inference model estimates the posterior probability that the current trial is a stop trial, $\sigma_t^{\nu} := P(s^{\nu} = 1 \mid y_{1:t}^{\nu})$. To compute this efficiently, the agent tracks an intermediate belief $\zeta_t^{\nu} := P(\theta^{\nu} \leq t \mid y_{1:t}^{\nu})$, representing the cumulative probability that the stop signal has *already* occurred.

To align with participants' behavioral constraints, we introduce three simplifying assumptions: (1) the stop belief is updated solely based on stop observations y_t^{ν}; (2) participants expect the stop signal onset based on a constant hazard rate λ; and (3) participants model the transient stop signal as persisting until the trial ends (although the input statistics change according to the bottom row of Table 2). This avoids complex inference about signal offset. It is justified empirically: only 10.03% of stop errors occurred during the post-signal fixation phase.

Based on these assumptions, the belief ζ_t^{ν} is updated recursively via Bayes rule. First, the predicted belief $\tilde{\zeta}_t^{\nu}$ incorporates the hazard function for the occurrence of the stop signal, h_t^{ν}, which combines the temporal prior λ and the stop prior $\tilde{r}^{\nu}$:

$$\tilde{\zeta}_t^{\nu} = \zeta_{t-1}^{\nu} + (1 - \zeta_{t-1}^{\nu}) \cdot h_t^{\nu} \tag{1}$$

$$h_t^{\nu} = \frac{\tilde{r}^{\nu}(1-\lambda)^{t-1}\lambda}{\tilde{r}^{\nu}(1-\lambda)^{t-1} + (1 - \tilde{r}^{\nu})} \tag{2}$$

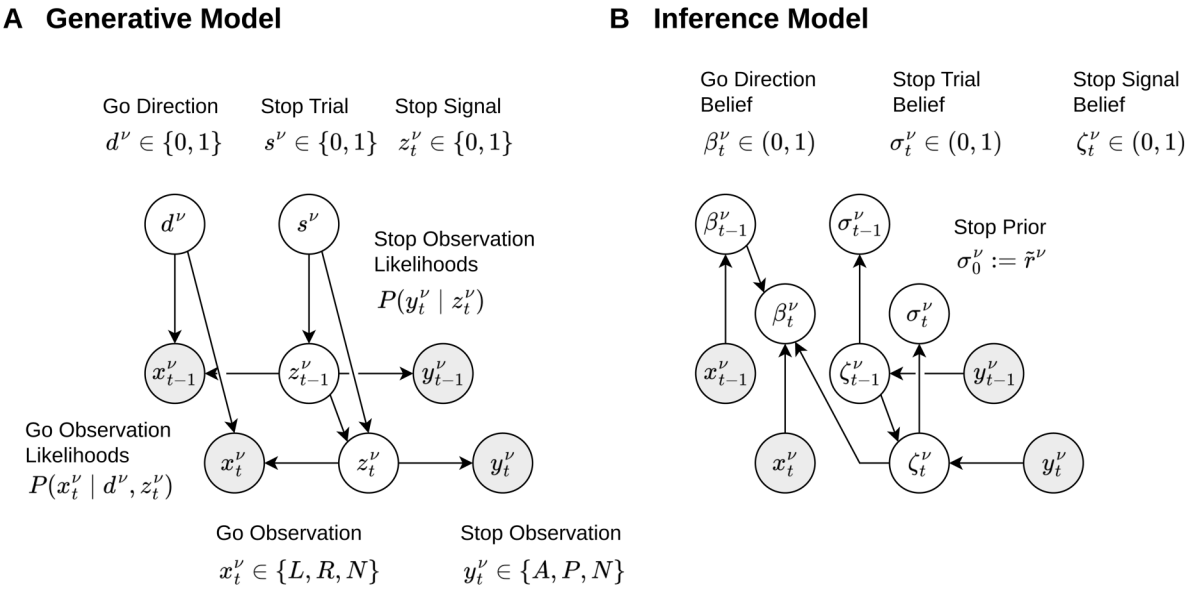


Fig 2. Generative and inference processes in the POMDP model. (A) Generative model. Each trial ν has a true go direction d^ν (Left or Right) and a true trial type s^ν (Go or Stop). At each time step t, s^ν and the Stop Signal Delay (SSD) determine the true stop signal state z_t^ν (Absence or Presence). The latent states d^ν, s^ν, and z_t^ν are hidden from the participant. Observable go cue inputs x_t^ν depend on both d^ν and z_t^ν. Observable stop signal inputs y_t^ν depend only on z_t^ν. **(B)** Inference model. At each time step t, the participant uses Bayes rule to calculate belief states (probabilities of the latent states). The stop signal state ζ_t^ν is inferred from y_t^ν and a hazard function. The trial type belief σ_t^ν is derived directly from ζ_t^ν. The go direction belief β_t^ν is inferred from x_t^ν and ζ_t^ν.

https://doi.org/10.1371/journal.pcbi.1014068.g002

Table 2. Observation likelihoods for the stop process.

Trial type s^ν	Time t	State z_t^ν	Observation likelihoods $P(y_t^\nu = \cdot \mid z_t^\nu)$		
			Absence *A*	**Presence *P***	**Null *N***
Go trial	$\forall t$	0	$(1-\delta')\delta$	$(1-\delta')(1-\delta)$	δ'
Stop trial	$t < \theta^\nu$	0	$(1-\delta')\delta$	$(1-\delta')(1-\delta)$	δ'
	$\theta^\nu \le t \le \theta^\nu + \mu$	1	$(1-\delta')(1-\delta)$	$(1-\delta')\delta$	δ'
	$t > \theta^\nu + \mu$	0	$\delta'\delta$	$\delta'(1-\delta)$	$1-\delta'$

θ^ν: stop signal onset time; μ: stop signal duration; $z_t^\nu = 1$ indicates stop signal presence (masking).

https://doi.org/10.1371/journal.pcbi.1014068.t002

Next, the posterior ζ_t^ν is corrected by the current sensory observation y_t^ν:

$$\zeta_t^\nu = \frac{P(y_t^\nu \mid z_t^\nu = 1) \cdot \tilde{\zeta}_t^\nu}{P(y_t^\nu \mid z_t^\nu = 1) \cdot \tilde{\zeta}_t^\nu + P(y_t^\nu \mid z_t^\nu = 0) \cdot (1 - \tilde{\zeta}_t^\nu)} \tag{3}$$

Finally, the total belief that the trial is a stop trial, σ_t^ν, integrates ζ_t^ν with the residual probability of a future onset:

$$\sigma_t^\nu = \zeta_t^\nu + (1 - \zeta_t^\nu) \cdot \frac{(1-\lambda)^t \cdot \tilde{r}^\nu}{(1-\lambda)^t \cdot \tilde{r}^\nu + (1 - \tilde{r}^\nu)} \tag{4}$$

Detailed derivation of stop process inference is in Section A.1 of the S1 Appendix.

The go process: generative and inference models In the generative model, the latent state $d^\nu \in \{0, 1\}$ denotes the true direction of the go cue (0 for Left, 1 for Right). Participants receive a noisy sensory input $x_t^\nu \in \{L, R, N\}$ (Left, Right, or Null) at each time step. This input is governed by the go precision (χ) and go null (χ') parameters (Table 3). On go trials, the cue remains visible throughout. On stop trials, the appearance of the stop signal at time θ^ν visually masks the go cue. This masking shifts the observation x_t^ν predominantly toward the Null status for the remainder of the trial (albeit allowing for a small residual signal associated with the go direction).

The inference model estimates the directional belief state $\beta_t^\nu := P(d^\nu = 1 \mid x_{1:t}^\nu, y_{1:t}^\nu)$. It is initialized uniformly at $\beta_0^\nu = 0.5$ (the probability of a left cue is $1 - \beta_t^\nu$). The presence of the stop signal ($z_t^\nu = 1$) masks the go cue and alters its observation likelihoods. Therefore, the inference for d^ν depends partially on the current stop signal belief ζ_t^ν.

The belief β_t^ν is updated recursively via Bayes rule by marginalizing over the stop signal state z:

$$\beta_t^\nu = \frac{\beta_{t-1}^\nu \cdot P(x_t^\nu \mid 1)}{\beta_{t-1}^\nu \cdot P(x_t^\nu \mid 1) + (1 - \beta_{t-1}^\nu) \cdot P(x_t^\nu \mid 0)} \tag{5}$$

$$P(x_t^\nu \mid d^\nu) = \underbrace{(1 - \zeta_t^\nu)}_{P(z_t^\nu = 0)} P(x_t^\nu \mid d^\nu, z_t^\nu = 0) + \underbrace{\zeta_t^\nu}_{P(z_t^\nu = 1)} P(x_t^\nu \mid d^\nu, z_t^\nu = 1) \tag{6}$$

where $P(x_t^\nu \mid d^\nu)$ is the marginal likelihood of observing x_t^ν given direction d^ν. An ambiguous sensory input ($x_t^\nu = N$) provides no directional evidence, resulting in no belief update ($\beta_t^\nu = \beta_{t-1}^\nu$).

The complete belief state at time t of trial ν is formalized as $\mathbf{b}_t^\nu = (\beta_t^\nu, \zeta_t^\nu)$. This joint state encodes the participant's inferred probabilities regarding both the go direction and the stop signal state. The trial type probability σ_t^ν is deterministically derived from ζ_t^ν (Eq 4). Therefore, maintaining ζ_t^ν acts as a sufficient statistic for the stop process within the state representation.

Optimal control. On each time step t, the participant chooses an action $a_t \in \{L, R, W\}$ (Go-Left, Go-Right or Wait). We formulate the action-value $Q_t(\mathbf{b}_t, a_t)$ in terms of expected *cost*, which is determined by the current belief state $\mathbf{b}_t = (\beta_t, \zeta_t)$ and intrinsic cost parameters. The objective is to minimize the total cost (Fig 3).

If the participant chooses a terminal response Go-Left (L) or Go-Right (R) at time $t \leq T$, the immediate cost incorporates three components: the accumulated time ($c_t \cdot t$), the expected directional error on go trials (c_{ge}), and the expected stop error on stop trials (c_{se}):

$$Q_t(\mathbf{b}_t, L) = c_t \cdot t + c_{\text{ge}} \cdot (1 - \sigma_t) \cdot \beta_t + c_{\text{se}} \cdot \sigma_t \tag{7}$$

Table 3. Observation likelihoods for the go process.

Trial type s^ν	Time t	State z_t^ν	State d^ν	Observation likelihoods $P(x_t^\nu = \cdot \mid d^\nu, z_t^\nu)$		
				Left L	Right R	Null N
Go trial	$\forall t$	0	0	$(1-\chi')\chi$	$(1-\chi')(1-\chi)$	χ'
			1	$(1-\chi')(1-\chi)$	$(1-\chi')\chi$	χ'
Stop trial	$t < \theta^\nu$	0	0	$(1-\chi')\chi$	$(1-\chi')(1-\chi)$	χ'
			1	$(1-\chi')(1-\chi)$	$(1-\chi')\chi$	χ'
	$t \geq \theta^\nu$	1	0	$\chi'\chi$	$\chi'(1-\chi)$	$1-\chi'$
			1	$\chi'(1-\chi)$	$\chi'\chi$	$1-\chi'$

θ^ν: stop signal onset time; μ: stop signal duration; $z_t^\nu = 1$ indicates stop signal presence (masking).

https://doi.org/10.1371/journal.pcbi.1014068.t003

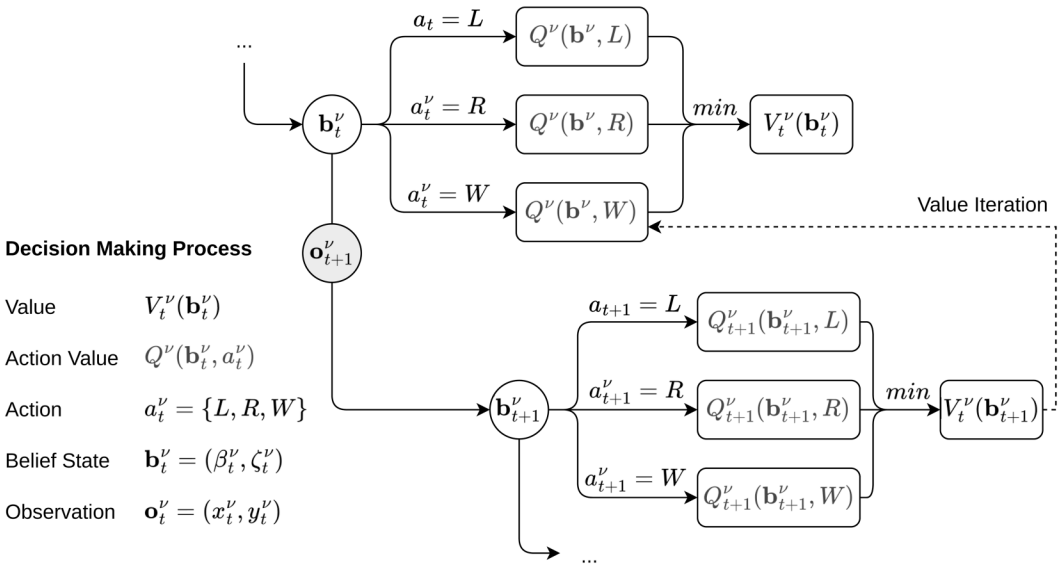


Fig 3. Optimal decision-making process in the POMDP model. At each time step t of trial ν, the belief state $\mathbf{b}_t^\nu = (\beta_t^\nu, \zeta_t^\nu)$ is updated using incoming go and stop observations (x_t^ν, y_t^ν). Action-values $Q_t^\nu(\mathbf{b}_t^\nu, a_t^\nu)$ represent the expected long-run cost of each action. They are computed from the current belief state and intrinsic costs. The agent selects the action from Go-Left ***(L)***, Go-Right ***(R)***, and Wait (*W*) that minimizes this expected cost. The value of the Wait action is derived recursively via value iteration (dotted line). This applies the Bellman equation to back-propagate expected future costs from V_{t+1}^ν to Q_t^ν.

https://doi.org/10.1371/journal.pcbi.1014068.g003

$$Q_t(\mathbf{b}_t, R) = c_t \cdot t + c_{\text{ge}} \cdot (1 - \sigma_t) \cdot (1 - \beta_t) + c_{\text{se}} \cdot \sigma_t \tag{8}$$

where the stop trial probability σ_t is derived directly from ζ_t.

Alternatively, the participant can choose to Wait (*W*) to accumulate more information. If *W* is chosen consistently until the terminal step *T*, a timeout occurs. This incurs time costs and a go missing penalty (c_{gm}):

$$Q_T(\mathbf{b}_T, W) = c_t \cdot T + c_{\text{gm}} \cdot (1 - \sigma_T) \tag{9}$$

For any $t < T$, the value of waiting is derived via the Bellman optimality equation [33]. It is calculated from the expected optimal value of the next belief state, marginalized over all possible observations at the next timestep $\mathbf{o}_{t+1} = (x_{t+1}, y_{t+1})$:

$$Q_t(\mathbf{b}_t, W) = \sum_{x_{t+1}} \sum_{y_{t+1}} P(x_{t+1}, y_{t+1} \mid \mathbf{b}_t, W) \cdot V_{t+1}^*(\mathbf{b}_{t+1}) \tag{10}$$

Here, $V_{t+1}^*(\mathbf{b}_{t+1}) = \min_a Q_{t+1}(\mathbf{b}_{t+1}, a)$ denotes the value of following the optimal policy from time $t+1$ onward. The transition probabilities $P(x_{t+1}, y_{t+1} \mid \mathbf{b}_t, W)$ are computed by marginalizing over the future stop signal state z_{t+1} and the latent go direction *d*. The detailed expansion of these transition probabilities is provided in Section A.3 of S1 Appendix.

The complete set of action-values is computed recursively via backward induction from *T* down to $t = 1$. For computational implementation, the continuous belief space is discretized into an 80×80 grid using a tensor-based approach (see Section A.4 of S1 Appendix).

Finally, to capture the stochasticity of human behavior, we model the actual action selection using a softmax policy applied to the pre-computed optimal costs:

$$\pi_t(a_t \mid \mathbf{b}_t) = \frac{\exp\left(-\varphi \cdot Q_t(\mathbf{b}_t, a_t)\right)}{\sum_{a_t \in \{L,R,W\}} \exp\left(-\varphi \cdot Q_t(\mathbf{b}_t, a_t)\right)} \tag{11}$$

where the inverse temperature parameter φ controls the degree of determinism in the response. The negative sign ensures that lower expected costs generate higher choice probabilities.

Estimation framework: TeSBI

Calculating the exact likelihood for the full POMDP model is computationally intractable due to the high-dimensional parameter space and the non-linear, discrete nature of the belief and action-value updates. Furthermore, observed behavioral sequences contain rich, multi-faceted temporal interactions (e.g., those induced by the SSD staircase) that are difficult to capture with simple summary statistics.

To address these challenges, we introduce Transformer-encoded Simulation-Based Inference (TeSBI). This approach extends standard SBI frameworks [25] by integrating a Transformer encoder [26] to learn compact, informative representations of the behavioral data. As illustrated in Fig 4, the TeSBI procedure consists of two stages: (1) Amortized training: mapping simulated behavioral sequences into embeddings via a Transformer encoder, and using a Neural Spline Flow (NSF) [24,34] to map these embeddings to parameter densities; and (2) Posterior inference: sampling individual-level posterior parameters from observed behavioral sequences using the trained network.

We established model reliability through parameter recovery (verifying the identifiability of all parameters from synthetic data) and posterior predictive checks (PPCs, confirming that simulated behavior generated from inferred parameters matches empirical patterns). Detailed algorithmic procedures, network architectures, and hyperparameter configurations, and transformer encoded representations are provided in Section B of S1 Appendix.

Amortized training. Feature engineering. To avoid the loss of information inherent to handcrafted summary statistics [35], we extract both trial-level and inter-trial features to prepare the full behavioral sequences as structured inputs for the downstream Transformer encoder. These arise from simulated sessions.

Each simulated session involves $N_{tr} = 360$ trials. For a given trial ν, the simulator produces a measurement tuple:

$$\mathbf{m}^{\nu} = (\mathrm{OUT}^{\nu}, \mathrm{RT}^{\nu}; s^{\nu}, \theta^{\nu})$$

where s^{ν} is the trial type, θ^{ν} is the SSD on stop trials (*null* on go trials), and $\mathrm{OUT}^{\nu} \in \{\mathrm{GS}, \mathrm{GE}, \mathrm{GM}, \mathrm{SS}, \mathrm{SE}\}$ denotes the behavioral outcome categories (go success, go error, go missing, stop success, stop error). As for the adaptive staircase

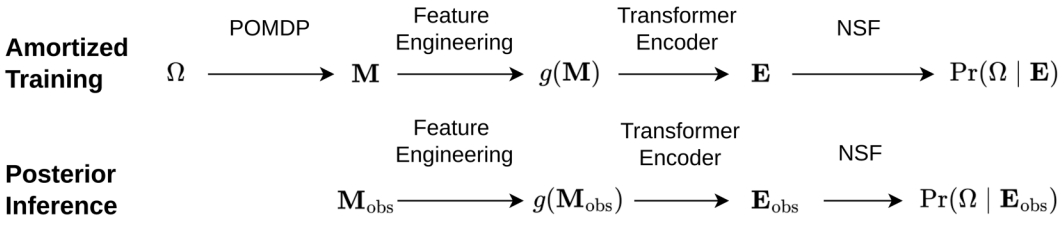


Fig 4. Overview of the end-to-end TeSBI procedure. (1) Amortized training: Parameters (Ω) sampled from the prior generate simulated behavior ($\mathbf{M}$) via the POMDP simulator. After feature engineering, a Transformer encoder extracts temporal embeddings ($\mathbf{E}$), and a Neural Spline Flow (NSF) is trained to map these embeddings to posterior densities. (2) Posterior inference: For each individual, the trained Transformer embeds their engineered observed sequence ($\mathbf{M}_{\mathrm{obs}}$), and the NSF estimates the parameter posteriors.

https://doi.org/10.1371/journal.pcbi.1014068.g004

procedure used in the original ABCD study, the SSD increases or decreases by 2 time steps (i.e., 50 ms) following a stop success or stop error, respectively, and is restricted to the range of [2,34] time steps (i.e., [50 ms, 850 ms]).

We denote the full session sequence as $\mathbf{M} = \{\mathbf{m}^{\nu}\}_{\nu=1}^{N_{\text{tr}}}$. To capture temporal dynamics, we employ a feature engineering operator $g(\cdot)$ that augments each trial into an 18-dimensional vector:

$$g(\mathbf{m}^{\nu}) = \big[\,\text{OneHot}(\text{OUT}^{\nu}), \text{RT}^{\nu}, \theta^{\nu}, \text{OneHot}(\text{OUT}^{\nu-1}), \text{RT}^{\nu-1}, \theta^{\nu-1}, \\ \mathbb{1}(\text{RT}^{\nu} = \varnothing), \mathbb{1}(\theta^{\nu} = \varnothing), \Delta\theta^{\nu}, \tilde{\nu}\big]$$

where OneHot$(\cdot)$ denotes a one-hot encoding of the 5 outcome types, and $\mathbb{1}(\cdot)$ is an indicator function checking for missing values ($\varnothing$). The vector also incorporates the previous trial's features, the change in SSD step-size ($\Delta\theta^{\nu} = \theta^{\nu} - \theta^{\nu-1}$), and a normalized trial index $\tilde{\nu} \in [-1, 1]$ tracking the trial's relative position within the session. For the first trial, previous features are initialized with zeros. Applying this operator to the full session yields the finalized sequence $g(\mathbf{M}) = \{g(\mathbf{m}^{\nu})\}_{\nu=1}^{N_{\text{tr}}}$, which is ready for downstream embedding.

Transformer encoder. To process these structured sequences, we employ a Transformer encoder f_{ϕ} as a feature extractor. Conceptually, we treat each trial as a *token* that contains information about the current behavior and context. The entire behavioral sequence acts as a *document*, the learned behavioral embedding serves as its *semantics*, and the underlying generative model parameters represent its core *themes*.

Formally, the encoder uses a form of self-supervision to map the transformed behavioral sequence into a dense summary embedding:

$$\mathbf{E} = f_{\phi}(g(\mathbf{M}))$$

from which it is ultimately possible to recover the posterior over the parameters that generated the sequence. Through its multi-head attention mechanism, the encoder naturally captures complex temporal structures (e.g., SSD adaptation dynamics, post-error slowing, preparatory effects, and fatigue) without requiring manual specification. The dense embedding $\mathbf{E}$ then serves as the direct input for the downstream density estimator.

Neural Spline Flow. We approximate the posterior distribution $P(\Omega \mid \mathbf{E})$ using a Neural Spline Flow (NSF) density estimator $q_{\psi}(\Omega \mid \mathbf{E})$ [34]. Conditioned directly on the summary embedding $\mathbf{E}$ extracted by the Transformer, the NSF and the encoder are optimized jointly across a large offline dataset of pre-simulated behavior. By minimizing the negative log-likelihood of the ground-truth generative parameters given these embeddings, the unified network learns a global mapping from complex behavioral patterns to parameter probability densities. The density estimation pipeline was implemented using the `zuko` normalizing flow library [36].

Posterior inference. As an amortized estimator, the fully trained TeSBI model performs zero-shot inference for all participants via a single forward pass. This eliminates the need for computationally expensive, subject-specific simulations.

For each participant u, the network directly maps their observed sequence $\mathbf{M}_{\text{obs}}^{(u)}$ to a summary embedding $\mathbf{E}_{\text{obs}}^{(u)}$. This embedding then conditions the NSF to estimate the individualized posterior:

$$\Omega^{(u)} \sim \Pr\left(\Omega \mid \mathbf{E}_{\text{obs}}^{(u)}\right)$$

We use the posterior mean as the point estimate for subsequent analyses, computed via 1,000 samples per participant.

Results

We characterize the behavior of the proposed POMDP framework and its application to the ABCD dataset. First, we illustrate the model's theoretical properties. This includes the within-trial dynamics of perceptual inference, the derived optimal policy, and the resulting action-values. Next, we validate the reliability of the TeSBI fitting pipeline. Finally, we present the empirical results, focusing on the distribution of estimated parameters and their association with ADHD scores.

Model simulations

To illustrate the internal dynamics of the proposed framework, we simulated behavior using our primary model (M6.1). Details regarding the selection of M6.1 (including model specification, parameter recovery, and posterior predictive checks) are provided in Section C of S1 Appendix.

Perceptual inference: Belief states dynamics. The temporal evolution of the agent's belief states reveals distinct patterns across five different trial outcomes (Fig 5). Here, we visualize example and summary simulations where the true go cue is Right and, on stop trials, the stop signal appears at SSD = 11.

In Go Success trials, β_t^ν accumulates towards 1, reflecting successful direction recognition (the white area indicates the proportion of trials where the model had already chosen to go). Conversely, in Go Error trials, β_t^ν drifts below 0.5,

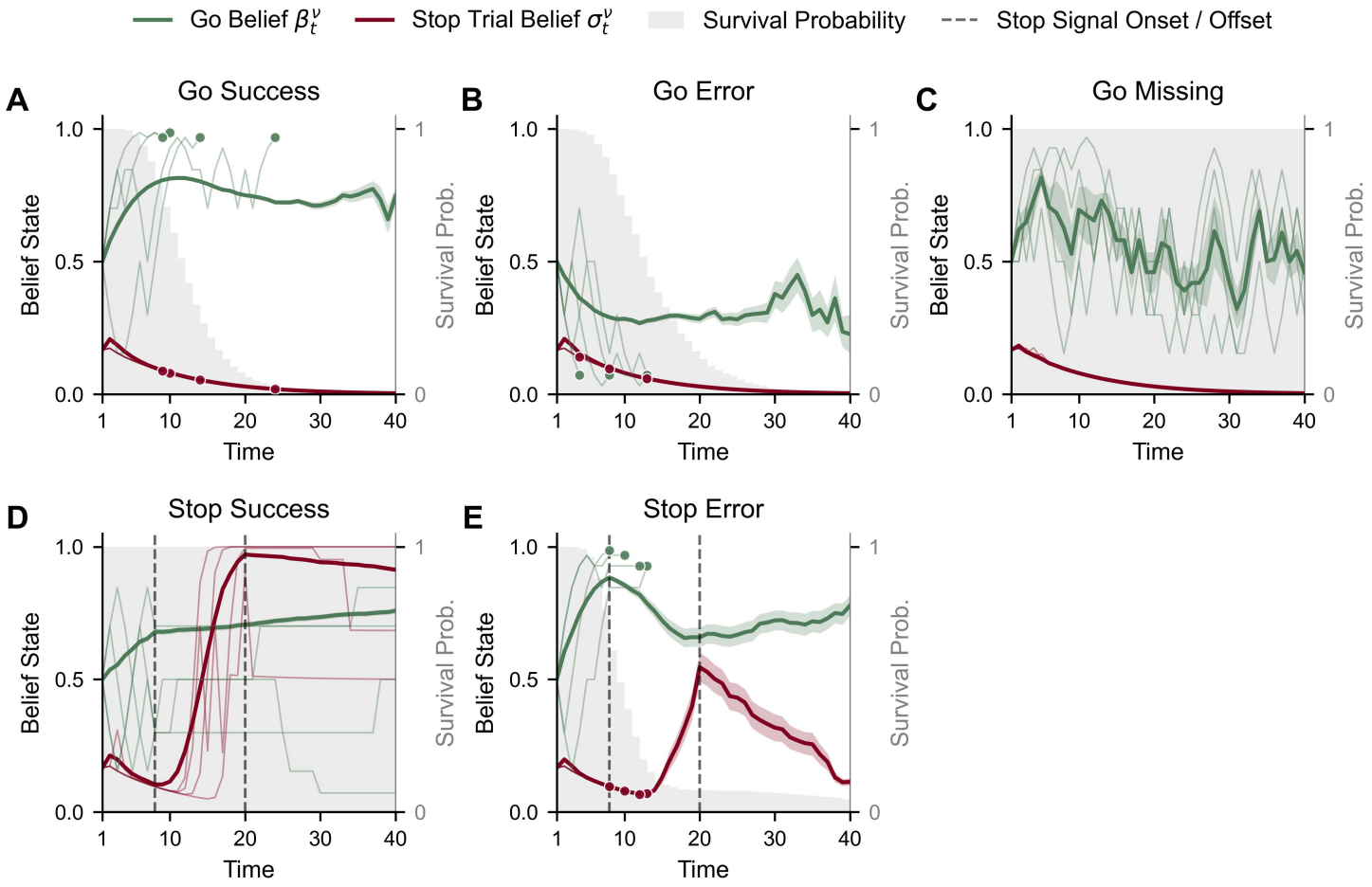


Fig 5. Belief state dynamics across five outcomes. Top row: go trials (with a right go cue). Bottom row: stop trials (with a right go cue and stop signal). Vertical dashed lines in the bottom panels indicate the stop signal window: onset at SSD = 11 (200 ms) and offset after a duration of 12 time steps (300 ms). Thin lines depict four randomly sampled individual trajectories per outcome; solid dots mark the decision times (RTs). Thick curves represent the mean belief trajectory ($\beta_t^\nu, \sigma_t^\nu$) computed over ongoing trials at each time step (shaded bands: s.e.m.; 6,000 go trials, 1,200 stop trials). Note that mean curves may not reach the trial horizon because individual trajectories are truncated at the go-decision point. The gray shaded area depicts the survival probability, indicating the proportion of trials that have not yet ended. Parameters are from a representative participant (the 95th percentile goodness-of-fit) derived from our primary model (M6.1): $\chi' = 0.080$, $\chi = 0.701$, $\delta' = 0.079$, $\delta = 0.901$, $c_{se} = 4.705$, $\varphi = 44.653$, $c_{ge} = 1.647$, $c_{gm} = 15.782$, $c_t = 0.046$, and $\tilde{r}^\nu = 0.167$. Although the fitted non-decision time is $\tau = 7$ steps (175 ms) for this subject, it is set to 0 here to visualize the cognitive process starting from $t = 1$.

https://doi.org/10.1371/journal.pcbi.1014068.g005

indicating that the agent erroneously accumulated observations for the left direction despite the cue being right. In Go Missing trials, β_t^ν stagnates around 0.5, implying that the agent fails to resolve directional ambiguity. This prompts persistent waiting until the time limit. Across all go trials, the stop signal belief (σ_t^ν, red line) consistently decays towards zero.

In Stop Success trials (successful inhibition), σ_t^ν initially decays but surges sharply immediately after the stop signal onset (first vertical dashed line), accurately tracking the signal's presence. It peaks before the signal offset and declines thereafter. In contrast, in Stop Error trials, the rise in σ_t^ν is notably delayed and attenuated. Simultaneously, the go belief β_t^ν often rises rapidly, overpowering the weaker stop belief and leading to a premature go response before the stop signal is fully processed.

The dynamics of β_t^ν and σ_t^ν reflect how sensory evidence for the go cue and stop signal accumulates over time. This establishes the perceptual basis for the subsequent optimal control.

Optimal control: Probabilistic policy and action values. Probabilistic policy. Fig 6 illustrates the POMDP choice strategy and belief state evolution. A softmax function maps *Q*-values across the joint belief space go belief (β_t^ν) and stop signal belief (ζ_t^ν) into regions favoring Wait (light gray), Go-Left (yellow), or Go-Right (blue). As the deadline ($t = 40$) approaches, the go regions expand inward, reflecting an increased urgency to act as the opportunity cost of a timeout outweighs the risk of an error. The full policy evolution is detailed in Section A.5 of S1 Appendix.

Realized actions depend on how belief states traverse this probabilistic space. In simulated trials (right-cue, SSD = 11), go success and go error trajectories rapidly accumulate directional evidence along the *y*-axis, triggering early responses in the Go-Right or Go-Left regions, respectively. Rare go missing trials occur when weak evidence leaves the state stranded in the central Wait region.

On stop trials, successful inhibition (stop success) occurs when signal detection rapidly increases ζ_t^ν, shifting the belief state rightward into the Wait region. Conversely, a stop error happens if go evidence accumulates too quickly, crossing the go threshold before the stop belief can sufficiently rise. Overall, these latent dynamics mutually corroborate the belief trajectories and action-values shown in Figs 5 and 7.

Action-values. Action-values (*Q*-values) encode expected long-run costs. Under the stochastic softmax policy, lower *Q*-values correspond to higher selection probabilities. Simulated *Q*-values are shown in Fig 7, following the same conditions as in Fig 5.

In Go Success trials, the *Q*-value of Wait is initially dominant (i.e., lowest). As the directional belief becomes sufficiently certain, the *Q*-value of the correct action (e.g., Go-Right) decreases, narrowing the gap with that of Wait. This relative reduction rapidly shifts the softmax probability in favor of the action, triggering a Go-Right response. Importantly, because the policy is stochastic, actions can be triggered even before a strict crossover in values occurs, provided the probability mass has shifted sufficiently. Moreover, while this shift is sharp in individual simulations (thin lines), the averaged trajectories (thick curves) appear smoother due to temporal variability across trials.

In Go Error trials, the *Q*-value of the incorrect action (e.g., Go-Left) decreases erroneously due to noisy belief accumulation favoring the wrong direction, resulting in a directional error. In Go Missing trials, Wait consistently retains a significantly lower *Q*-value than the alternatives, keeping the probability of go actions negligible throughout the trial.

In Stop Success trials, Wait remains the optimal action (lowest *Q*-value) throughout the trial. Crucially, this sustained waiting reflects strategic inhibition under risk, not perceptual uncertainty. As shown in Fig 5, the agent quickly detects the stop signal (rapid rise in stop belief σ_t^ν); this high certainty, combined with the high cost of stop errors (c_{se}), drives the *Q*-value of Wait down to enforce inaction (rather than prolonged uncertain waiting).

In Stop Error trials, the *Q*-value of a go action becomes competitive with Wait despite the stop signal. Multiple interacting factors, such as stronger early go directional observations, weaker stop signal detection, or a lower stop error penalty, collectively shift the action-value to favor execution over inhibition.

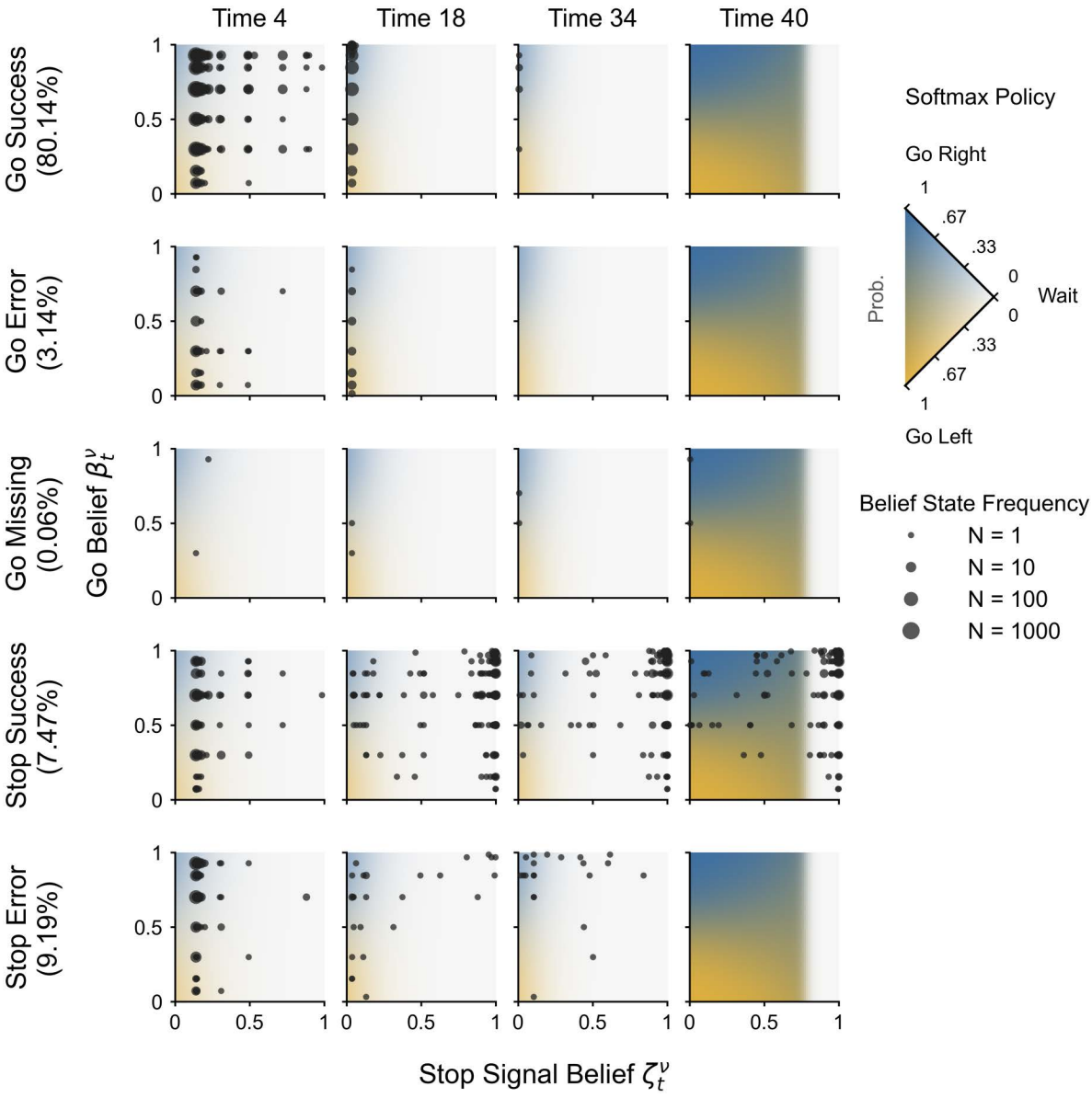


Fig 6. Softmax policy and belief state trajectories across selected time steps and trial outcomes. Background colors represent the softmax policy probabilities at selected time steps across the joint belief space: go belief (β_t^v) and stop signal belief (ζ_t^v). Because the softmax function applies a highly non-linear transformation that compresses extreme *Q*-values, the color mapping is visually enhanced using gamma correction (gamma is 0.05). The simplex legend on the right reflects this corrected mapping, indicating the probability distribution between Go-Right (blue), Go-Left (yellow), and Wait (light gray). Overlaid black dots represent the agent's belief states, drawn from simulations of 6,000 go trials (right go cue) and 1,200 stop trials (right go cue; stop signal with SSD = 11). The trials are categorized into five outcomes with their respective simulated frequencies. The size of the belief state dots represents their density. Trajectories terminate at the go-decision point, while states in Go Missing and Stop Success outcomes persist in the central Wait region until the deadline. Parameters are identical to those in Fig 5.

https://doi.org/10.1371/journal.pcbi.1014068.g006

Model selection and validation

To ensure parameter identifiability, we evaluate 15 candidate POMDP variants using an iterative step-down procedure. Starting from the unconstrained full model, we systematically reduce unidentifiable parameters by fixing those that fail the

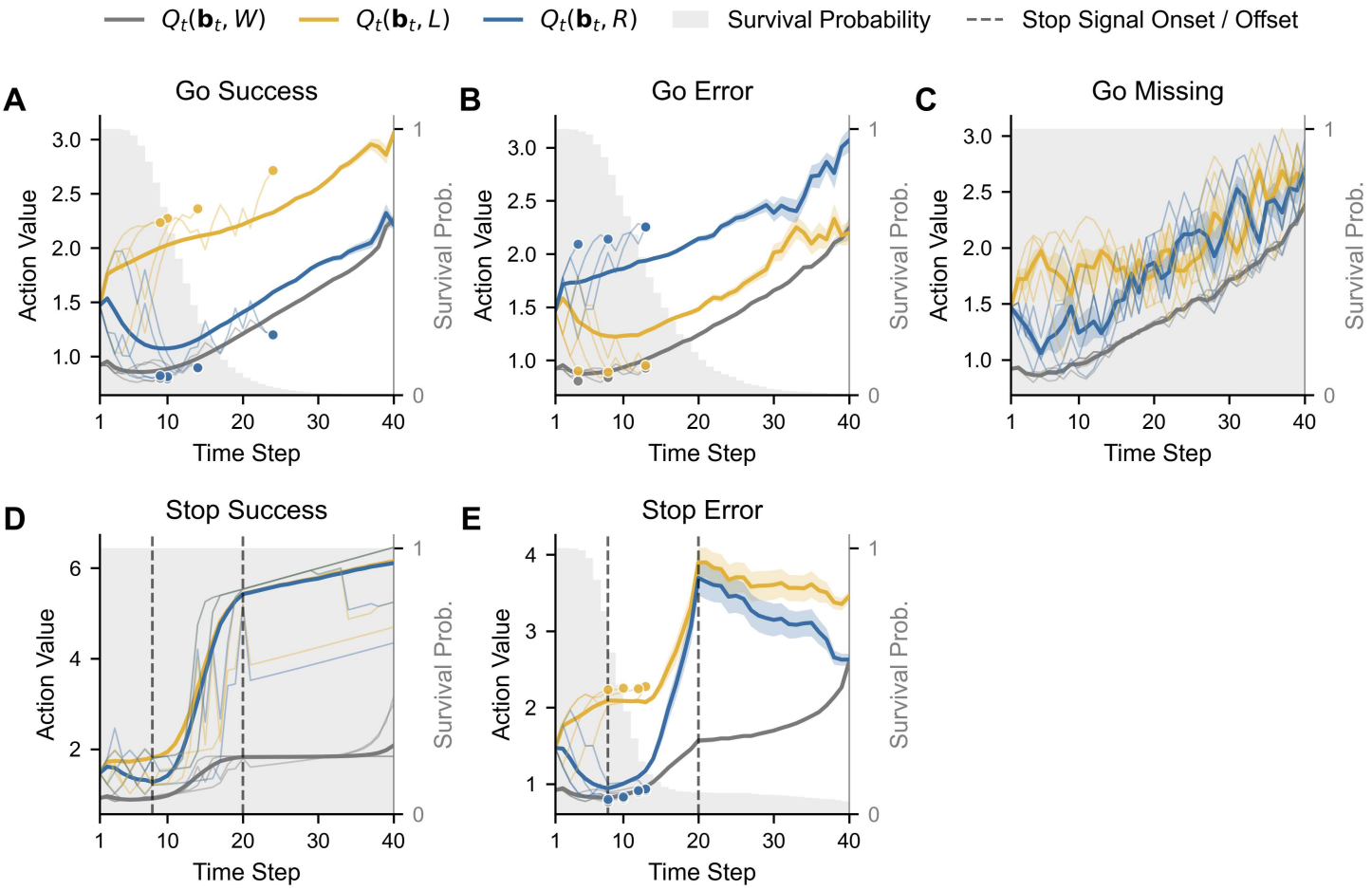


Fig 7. Action-value dynamics across five outcomes. Thick curves represent the mean action-value trajectory $Q_t^{\nu}(\cdot)$ for Go-Left, Go-Right, and Wait actions, computed over ongoing trials at each time step (shaded bands: s.e.m.; 6,000 go and 1,200 stop trials). A softmax policy is used for decision-making: actions with lower expected costs (*Q*-values) are optimal and more likely to be chosen, though stochasticity permits the occasional selection of higher-cost actions. Plotting conventions (trial layouts, stop signal window, sample trajectories, decision markers, and survival probability) and all model parameters are identical to those in Fig 5.

https://doi.org/10.1371/journal.pcbi.1014068.g007

recovery test (Pearson's $r < 0.65$) to specific posterior quantiles derived from the preceding model. We repeat this reduction process until all remaining free parameters are robustly identifiable.

We then evaluate the fit of these models using in-sample and out-of-sample posterior predictive checks (PPCs). We quantify the distance between behavioral observations and simulations via an aggregate distance metric comprising nine components: choice proportions (GS, GE, GM, SS), Wasserstein distances for RT distributions (GS, GE, SE), and Kolmogorov-Smirnov distances for RT distributions (GS, SE).

We select M6.1 as the primary model, as it yields the lowest total distance among the fully identifiable variants. It successfully recovers all six core parameters, showing strong correlations for go precision (χ, Pearson's $r = 0.89$), non-decision time (τ, $r = 0.83$), time cost (c_t, $r = 0.82$), and stop error cost (c_{se}, $r = 0.81$), alongside reliable recovery for go error cost (c_{ge}, $r = 0.73$), noting the rarity of go error behaviors; and stop precision (δ, $r = 0.73$), especially given the sparsity of stop trials in the task design.

To illustrate the predictive power of M6.1, Fig 8 shows fits for three participants representing Good, Moderate, and Poor fits (the 95th, 50th, and 5th percentiles of the aggregate distance, respectively). Across these performance levels, the model accurately reproduces the empirical outcomes, RT distributions (go and stop error), and task dynamics (SSD tracking and stop success rates). Even for the "Poor" fit participant, the model provides a reasonable approximation; the higher distance metric primarily reflects the participant's inherently noisy behavior (e.g., non-monotonic inhibition across the SSD staircase) rather than structural model failures.

Comprehensive details regarding the step-down model selection process, complete parameter recovery matrices, and PPCs results are provided in Section C of S1 Appendix.

Beyond individual parameter associations, we validate the holistic representations learned by the Transformer encoder using Canonical Correlation Analysis (CCA) and Principal Component Analysis (PCA). CCA reveals a strong, multidimensional correspondence between latent embeddings and model-agnostic behavioral metrics (primary variate $\rho_c = 0.995$, $p < 0.001$), confirming the encoding of both static and temporal task dynamics. Furthermore, PCA projections reveal that these behavioral fingerprints form a continuous computational manifold, confirming that higher ADHD scores are heterogeneously distributed across the population rather than clustered as an isolated deficit. Detailed cross-loadings and visualizations are provided in Section B.2 of S1 Appendix.

Computational attributes and ADHD score associations

Here, we define the six core continuous parameters derived from our POMDP model collectively as latent computational attributes. These attributes quantify distinct aspects of an individual's cognitive architecture, specifically capturing the characteristics of their perceptual processing and subjective valuation. To obtain these metrics, we draw 1,000 samples from the posterior parameter distribution of each participant and compute the mean as an individual-level point estimate. The aggregate parameter distributions across the valid baseline year sample are provided in Section E.1 of S1 Appendix.

To investigate how computational attributes relate to clinical characteristics, we analyzed the association between the six estimated POMDP parameters and ADHD scores. We included sex, intelligence quotient (IQ, derived from the WISC Matrix Reasoning score [37]), and medication status (coded as 1 for stimulant prescription [38]) as established covariates.

To facilitate comparability, all continuous variables, including the estimated POMDP parameters, ADHD scores, and IQ scores, were standardized (z-scored) prior to the analysis. We then regressed each parameter on these predictors using multiple linear regression. To ensure model parsimony, we performed stepwise model selection via the Bayesian Information Criterion (BIC). Because BIC consistently penalized the inclusion of interaction terms across all parameters, we retained the simplified main-effects model for all subsequent analyses (see Section E.2 of S1 Appendix for detailed comparisons).

Within the main-effects model (Table 4), higher ADHD scores significantly predict lower go precision (χ; $\beta = -0.068$, $p < .001$), lower stop error cost (c_{se}; $\beta = -0.053$, $p = .004$), lower time cost (c_t; $\beta = -0.050$, $p = .007$), and lower go error cost (c_{ge}; $\beta = -0.043$, $p = .020$). We observe no significant associations with stop precision (δ) or non-decision time (τ).

Covariates exhibit distinct computational profiles. Males show higher stop error (c_{se}) and time costs (c_t), alongside lower go error costs (c_{ge}) compared to females, with no significant differences in precision or non-decision time. Higher IQ robustly predicts higher go precision (χ) and longer non-decision time (τ), along with higher stop error and time costs, though it is associated with slightly lower stop precision (δ). Medication status shows significant positive associations with go precision and go error cost. Notably, while statistically robust, absolute effect sizes remain generally small (partial $R^2 < 0.015$, Cohen's $f^2 < 0.015$). However, ADHD characteristics and established covariates (Sex and IQ) exhibit comparable, parameter-specific effect size profiles.

Beyond mechanistic insights, POMDP parameters explain slightly more variance in ADHD scores than classic behavioral metrics after controlling for covariates. Adding these parameters to classic metrics significantly improves overall

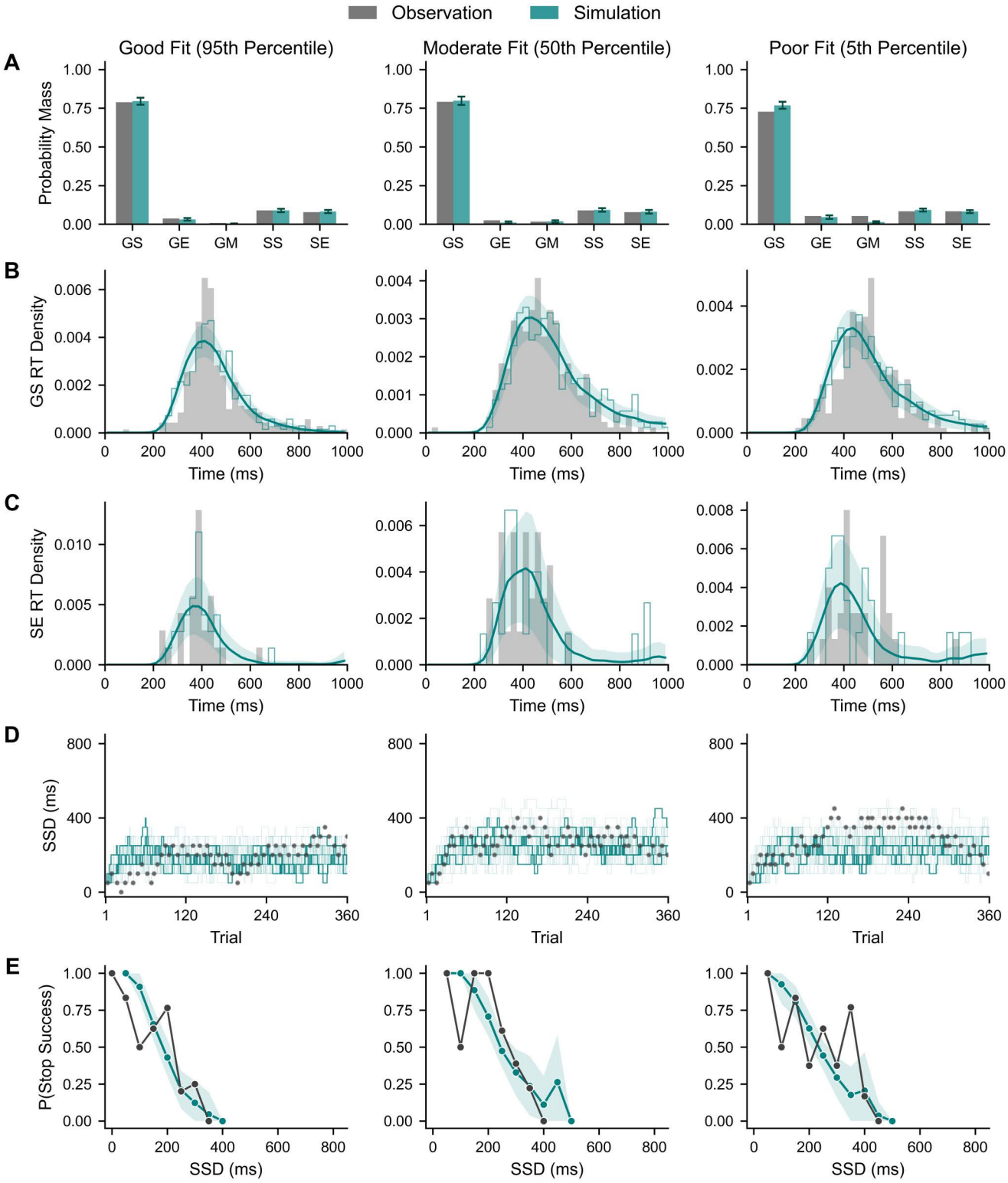


Fig 8. Representative posterior predictive checks for M6.1. Three participants illustrate Good (left), Moderate (middle), and Poor (right) model fits. **(A)** Observed outcome rates (gray) compared to the mean of 30 simulations (teal, generated using posterior mean parameters), with error bars

indicating s.e.m. across simulations. **(B–C)** Observed reaction time distributions for Go Success and Stop Error (gray) compared to the mean simulated density (teal) and its variability (shaded area s.e.m.). A single simulation trace (thin teal line) is overlaid to illustrate a typical simulation. **(D)** SSD staircase dynamics. Observed trials (gray) are plotted alongside all simulated trajectories (faint teal), with three individual simulations (teal) highlighting the model's adaptive behavior. **(E)** Inhibition dynamics. The observed probability of stop success (gray line and dots) is compared to the simulated mean (teal line and dots) and its variability (shaded area s.e.m.). All model time steps are converted to real time (1 time step = 25 ms).

https://doi.org/10.1371/journal.pcbi.1014068.g008

Table 4. Parameter Regression Coefficients and Effect Sizes.

Parameters	Coef.	SE	*p*-value	Partial R^2	f^2
ADHD					
χ	-0.0681	0.0182	0.0002***	0.0039	0.0039
δ	-0.0038	0.0184	0.8351	<0.0001	<0.0001
c_{se}	-0.0533	0.0183	0.0036**	0.0024	0.0024
c_{ge}	-0.0427	0.0184	0.0201*	0.0015	0.0015
c_t	-0.0498	0.0183	0.0067**	0.0021	0.0021
τ	-0.0256	0.0184	0.1635	0.0005	0.0005
Sex (Male)					
χ	-0.0244	0.0336	0.4684	0.0001	0.0001
δ	-0.0435	0.0340	0.2017	0.0005	0.0005
c_{se}	0.2143	0.0338	<0.0001***	0.0111	0.0113
c_{ge}	-0.1094	0.0340	0.0013**	0.0029	0.0029
c_t	0.1584	0.0339	<0.0001***	0.0061	0.0061
τ	0.0006	0.0340	0.9866	<0.0001	<0.0001
IQ					
χ	0.1455	0.0166	<0.0001***	0.0211	0.0216
δ	-0.0374	0.0168	0.0264*	0.0014	0.0014
c_{se}	0.0362	0.0167	0.0306*	0.0013	0.0013
c_{ge}	0.0242	0.0168	0.1496	0.0006	0.0006
c_t	0.0492	0.0168	0.0033**	0.0024	0.0024
τ	0.0534	0.0168	0.0015**	0.0028	0.0028
Medication					
χ	0.2782	0.0789	0.0004***	0.0035	0.0035
δ	-0.0793	0.0799	0.3210	0.0003	0.0003
c_{se}	0.1109	0.0794	0.1628	0.0005	0.0005
c_{ge}	0.1618	0.0798	0.0426*	0.0012	0.0012
c_t	0.0706	0.0796	0.3752	0.0002	0.0002
τ	-0.0655	0.0798	0.4122	0.0002	0.0002

Note: $N = 3{,}566$. * $p < .05$, ** $p < .01$, *** $p < .001$.

https://doi.org/10.1371/journal.pcbi.1014068.t004

model fit, demonstrating they capture unique clinical information missed by conventional summaries (see Section E.3 of S1 Appendix).

We can illustrate these relationships by visualizing the distributions of the computational attributes stratified by ADHD scores and sex (Fig 9). As expected from the regressions, we observe population-level shifts across the clinical scores. For instance, the parameters significantly associated with ADHD, go precision (χ), stop error cost (c_{se}), go error cost (c_{ge}), and time cost (c_t), highlighted with dashed mean lines, exhibit leftward shifts as ADHD scores increase. The distributions also highlight robust sex differences in specific cost valuations. Across all ADHD severity bins, males consistently show

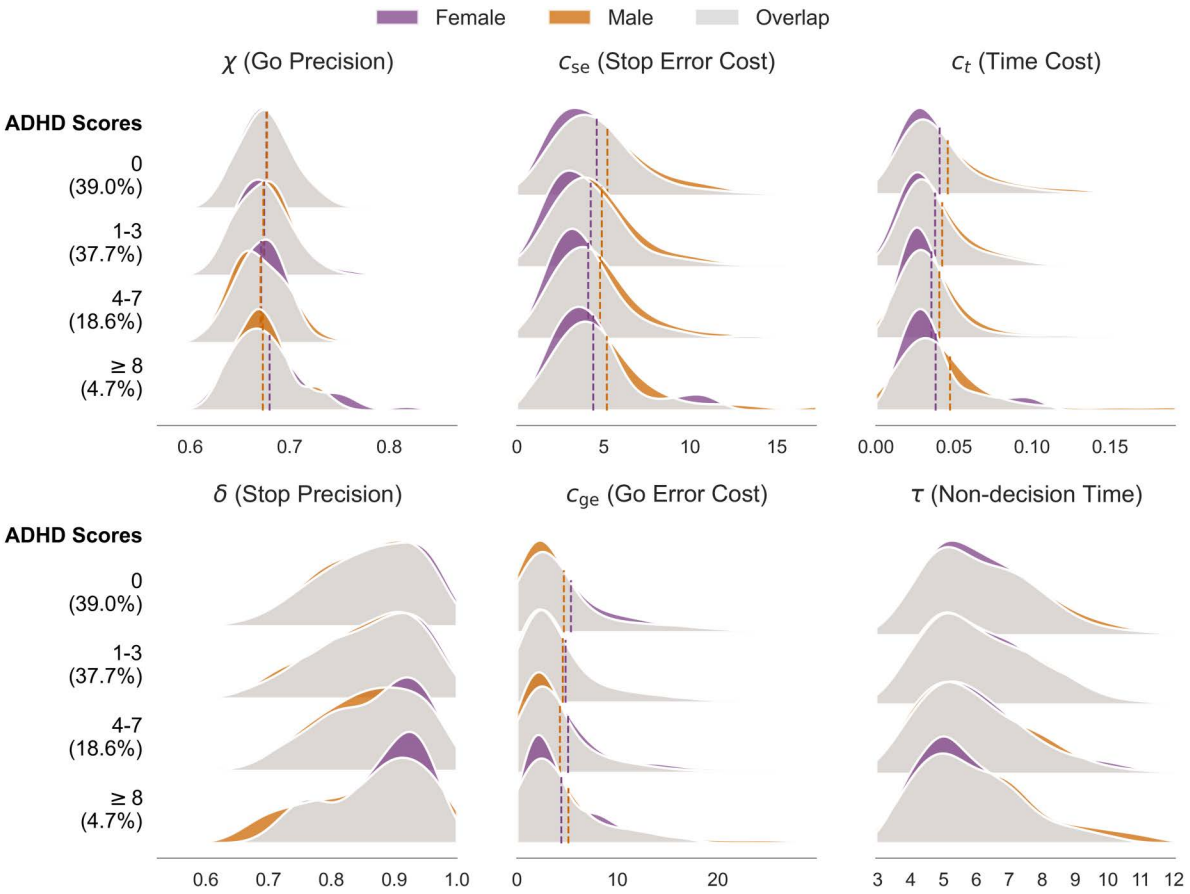


Fig 9. Distributions of participant-level computational attributes by ADHD scores and sex. Posterior distributions of the six POMDP parameters across the valid baseline sample ($N = 3{,}567$). Participants are categorized into four clinical bins based on their ADHD scores (0, 1-3, 4-7, ≥8). Density plots are colored by sex, with the overlapping intersection shaded in gray. Vertical dashed lines denote the group-specific means for the four parameters exhibiting significant main effects of ADHD. The visualization highlights three key features of the computational attributes: progressive mean shifts across the clinical gradient, with four parameters showing significant decreases as ADHD scores increase (go precision χ, $p < .001$; c_{se} and c_t, $p < .01$; c_{ge}, $p < .05$); robust sex differences confined to the three cost valuations (c_{se}, c_{ge}, c_t, all $p < .01$); and general individual heterogeneity, evidenced by wide probability densities and extensive overlap even between extreme clinical subgroups.

https://doi.org/10.1371/journal.pcbi.1014068.g009

higher c_{se} and c_t, and lower c_{ge} compared to females, maintaining a visible gap between their respective group means. However, although these shifts are highly statistically significant given the large sample size, they appear visually subtle due to the small absolute effect sizes.

Crucially, the plots reveal substantial within-group heterogeneity. The wide probability densities and the extensive overlap between distinct clinical and demographic subgroups emphasize that ADHD is computationally heterogeneous. Rather than forming a homogeneous cluster, individuals with identical clinical scores can possess vastly different underlying cognitive architectures.

In summary, the computational attributes of higher ADHD scores stem from degraded information processing and altered cost valuation. This involves a reduction in go cue directional precision, alongside a blunted sensitivity to both errors and time. Furthermore, the systematic exclusion of interaction terms demonstrates that these core computational alterations remain highly consistent regardless of patient sex, IQ, or medication status.

Discussion

Summary

We develop a Partially Observable Markov Decision Process (POMDP) model of the Stop Signal Task (SST) and apply it to the large-scale ABCD baseline dataset. Following previous work formalizing inhibitory control as a continuous process of Bayesian inference and value-based action [21], our model unifies perception and decision-making within a single framework. To overcome the computational bottleneck of fitting such complex models, we introduce Transformer-encoded Simulation-Based Inference (TeSBI), an end-to-end pipeline enabling efficient and reliable parameter estimation at scale.

Applying this pipeline to the ABCD cohort, we systematically reduce the number of free parameters to ensure identifiability via parameter recovery and posterior predictive checks, resulting in a robust six-parameter POMDP model. We obtain subject-level parameters as computational attributes, which are then regressed on ADHD scores while controlling for sex, IQ, and medication status. Our main-effects regression reveals that, on average, higher ADHD scores are most robustly associated with a reduction in go cue directional precision, alongside a blunted sensitivity to stop error, go error and time costs.

Crucially, while these regressions capture overarching group-level trends, our latent behavioral embeddings reveal that these computational attributes form a continuous spectrum rather than discrete clinical clusters. The broad dispersion of individuals with higher ADHD scores across this space highlights substantial heterogeneity in their cognitive strategies, nuances that simple group-level averages inherently overlook.

Mechanistic insights

The dominant accounts of the Stop Signal Task, initiated by the independent race model [4,5] and its interactive variants [13,14,39], have provided profound insights into the architecture of cognitive control. By conceptualizing inhibition as a mechanical race between Go and Stop processes, these frameworks established the foundational methodology of using summary statistics to estimate the unobservable Stop Signal Reaction Time (SSRT) [40].

Building upon this foundation, advanced parametric extensions have further refined our understanding. Models such as the Bayesian estimation of Ex-Gaussian SSRT distributions (BEESTS) [41,42], the racing diffusion Ex-Gaussian (RDEX-ABCD) model [43], the racing diffusion model (RDM) [44,45] (which integrates the drift diffusion model (DDM) [7,12,46,47]), and the linear ballistic accumulator (LBA) for the SST [48,49], offer greater precision by fitting the full distributions of reaction times and accuracy. These approaches demonstrate how the *accumulate-to-bound* framework can capture the temporal dynamics of finishing times. However, by focusing primarily on perceptual accumulation, they typically abstract away the subjective valuation process, where an agent continuously weighs the cost of delaying an action against the risk of making an error at each moment.

Our POMDP framework explicitly addresses this gap by extending the foundational conceptualization of Shenoy and Yu [22], formalizing how the well-documented *race-like* dynamics of inhibition can naturally emerge from dynamic belief updates and continuous value optimization. This conceptual shift provides a normative perspective on how subjective sensory uncertainty and intrinsic costs optimally shape these racing processes. This approach offers three key advantages.

First, regarding *perceptual inference*, we augment the original framework by introducing subjective sensory noise to explicitly capture perceptual ambiguity. Furthermore, we generalize the model structure to accommodate a dependent SST design, in which the stop signal masks the go cue. By modeling this interaction such that the agent's rising belief in the stop signal mathematically degrades its certainty regarding the go cue, the perceptual masking effect emerges as a natural property of Bayesian inference. Ultimately, this dependent framework subsumes the traditional independent processing assumption as a special case.

Second, regarding the *optimal policy*, our model relies on dynamic, value-based action valuation rather than fixed decision thresholds. While standard models such as the DDM constitute the asymptotically optimal policy for simple 2AFC

tasks [19,50,51], the continuous uncertainty and conflicting goals of the SST require the moment-by-moment, value-driven trade-off between the different possible trial types to be captured. From this perspective, reaction time slowing is interpreted as an active, strategic adjustment to minimize expected costs, rather than a passive braking mechanism. Section D of the S1 Appendix provides a qualitative comparison between our value-based mechanism and a more standard accumulate-to-bound architecture (RDEX-ABCD), a race model adapted to the peculiarities of the ABCD task [43]. The models fit comparably well, opening an important question for the future in mapping their rather different parameterizations onto each other.

Crucially, our computational perspective offers new clinical insights. While prior studies have highlighted stop-signal perceptual deficits as a plausible mechanism underlying ADHD [9], our findings tentatively suggest an alternative perspective (albeit with very small effect sizes). Within a unified framework like the POMDP, we hypothesize that a substantial portion of the behavioral variance traditionally attributed to pure perceptual impairments might also stem from top-down alterations in how individuals subjectively weigh the costs of errors and delays. Similarly, while recent DDM-based analyses of the ABCD task link ADHD to longer non-decision times [43], we observed only a non-significant negative trend. This divergence likely arises because our included covariates (sex, IQ, and medication) captured significant variance in non-decision time, combined with the limited clinical ADHD severity in the ABCD study.

Finally, regarding computational tractability, the introduction of *TeSBI* addresses the computational challenges that have typically restricted the application of complex cognitive models to large-scale cohorts. Unlike conventional summary-statistic methods, our Transformer-based encoder captures the sequential dependencies of trial-by-trial data. This enables the efficient and reliable estimation of individual-level posterior distributions, facilitating high-throughput computational phenotyping.

Limitations and future directions

Our study has several limitations that highlight important avenues for future research. These can be broadly categorized into modeling and inference constraints, data characteristics, and clinical applications.

First, regarding modeling constraints, we assume deterministic time perception (i.e., known trial duration T). In reality, however, human time perception is inherently noisy [52]. Incorporating stochastic timing (e.g., $\tilde{T} \sim \mathcal{N}(T, \sigma_T^2)$) would allow agents to act based on an approximation of the remaining time. This adjustment would smooth the value transitions near the deadline and better capture the variability of premature and late responses.

Second, our current generative model omits the dynamic anticipation of stop signals. Within a trial, we do not model the participant's prior expectation of exactly when the stop signal would appear (assuming a constant within-trial hazard rate λ). Across trials, we also abstract away how sequential history influences prior beliefs. While the SSD is primarily determined by a self-adaptive staircase procedure based on previous Stop outcomes, the number of consecutive Go trials also indirectly shapes behavior. Specifically, a long sequence of Go trials may initially decrease the predicted probability of an upcoming Stop trial (via Bayesian updating [20,22,53,54]) and subsequently increase it as participants expect Stop trials to recur. However, because the 1/6 stop ratio in the ABCD task restricts the fluctuation of this prior, fixing it to 1/6 allows us to focus on within-trial dynamics. As shown in our sensitivity analysis (Section A.6 of S1 Appendix), variations within this restricted range generate small but observable shifts in performance; thus, fixing the prior simplifies the current model. This approach establishes stable within-trial parameters first, allowing future extensions to free the stop prior and investigate across-trial effects.

Third, concerning inference constraints, we generate all simulated behaviors upfront before model training. Although we utilize parallel computing, matrix vectorization, and early stopping to accelerate the process, simulating data on CPUs remains a bottleneck compared to gradient-based training on GPUs. Future implementations could adopt an active learning or multi-round approach (i.e., iteratively generating data, training the model, and evaluating results) to further optimize computational resources, especially for models that feature dynamic learning across trials.

Regarding clinical relevance, we rely on ADHD scores assessed via the CBCL questionnaire rather than formal clinical diagnoses. The ABCD cohort is generally healthy, with relatively few participants exhibiting severe ADHD symptoms (see Section F of S1 Appendix for statistics). While this approximates the distribution in the general population, the lack of severe cases might lead to an underestimation of ADHD-related effects. Validating these computational attributes in clinical datasets with a larger proportion of diagnosed individuals is necessary to confirm the robustness of our results.

Furthermore, our current analysis is cross-sectional, relying exclusively on baseline data. Because the ABCD study features a longitudinal design, future work can track the developmental trajectories of these computational attributes and ADHD effects over a longer period (e.g., the ten-year span of the ABCD protocol). This offers a unique opportunity to investigate whether specific model parameters can serve as early predictors of symptom progression or as treatment biomarkers [55].

Finally, while this study focuses strictly on behavior, the availability of comprehensive neuroimaging data in the ABCD study is a tempting opportunity. Future research can link these computational parameters to specific neural correlates, potentially establishing them as quantitative endophenotypes that bridge the gap between underlying neural dysregulation and clinical behavior [56].

Conclusion

In total, we built a POMDP model able to accommodate the rather particular version of the stop signal task implemented in the ABCD study, alongside an inference pipeline capable of fitting participants at an appropriate scale. Along with illustrating some of the peculiarities of this version, we showed significant but subtle correlations between the fit parameters and the rather limited range of ADHD scores in the study.

Supporting information

S1 Appendix. Supplementary methods and results. This appendix provides comprehensive details supporting the main contents, including: (A) Mathematical details of the POMDP model, covering derivations of stop process inference, the formulation for independent stop signal tasks, transition probabilities, tensor-based value iteration, full probabilistic policy visualization, and stop prior sensitivity analysis; (B) TeSBI architecture and validation, detailing the base architecture and transformer-encoded representations; (C) POMDP model selection and validation, including model specification, model selection procedure, parameter recovery analysis, and posterior predictive checks; (D) Comparison with alternative models; (E) Computational attributes, including aggregate distributions, parameter regression and interaction analysis, and predictor sets comparison; and (F) Model-agnostic statistical analysis.
(PDF)

Acknowledgments

We thank Prof. Jakob Macke and Dr. Cornelius Schröder for their valuable advice on simulation-based inference. We also thank our colleagues from both laboratories for their support and the wider research community for inspiring discussions at conferences. We acknowledge the use of Gemini (Google) for assistance with LaTeX formatting and code correction.

Author contributions

Conceptualization: Wenting Wang, Tobias Kaufmann, Peter Dayan.

Data curation: Wenting Wang.

Formal analysis: Wenting Wang, Peter Dayan.

Funding acquisition: Tobias Kaufmann, Peter Dayan.

Methodology: Wenting Wang, Peter Dayan.

Project administration: Wenting Wang, Tobias Kaufmann, Peter Dayan.

Resources: Tobias Kaufmann, Peter Dayan.

Software: Wenting Wang.

Supervision: Tobias Kaufmann, Peter Dayan.

Validation: Wenting Wang, Tobias Kaufmann, Peter Dayan.

Visualization: Wenting Wang.

Writing – original draft: Wenting Wang, Peter Dayan.

Writing – review & editing: Wenting Wang, Tobias Kaufmann, Peter Dayan.

References

1. Barkley RA. Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. Psychol Bull. 1997;121(1):65–94. https://doi.org/10.1037/0033-2909.121.1.65 PMID: 9000892
2. Hill EL. Executive dysfunction in autism. Trends Cogn Sci. 2004;8(1):26–32. https://doi.org/10.1016/j.tics.2003.11.003 PMID: 14697400
3. Jalal B, Chamberlain SR, Sahakian BJ. Obsessive-compulsive disorder: Etiology, neuropathology, and cognitive dysfunction. Brain Behav. 2023;13(6):e3000. https://doi.org/10.1002/brb3.3000 PMID: 37137502
4. Logan GD, Cowan WB, Davis KA. On the ability to inhibit simple and choice reaction time responses: a model and a method. J Exp Psychol Hum Percept Perform. 1984;10(2):276–91. https://doi.org/10.1037/0096-1523.10.2.276 PMID: 6232345
5. Logan GD, Cowan WB. On the ability to inhibit thought and action: A theory of an act of control. Psychol Rev. 1984;91(3):295–327. https://doi.org/10.1037/0033-295x.91.3.295
6. Verbruggen F, Logan GD. Response inhibition in the stop-signal paradigm. Trends Cogn Sci. 2008;12(11):418–24. https://doi.org/10.1016/j.tics.2008.07.005 PMID: 18799345
7. Ratcliff R, McKoon G. The diffusion decision model: theory and data for two-choice decision tasks. Neural Comput. 2008;20(4):873–922. https://doi.org/10.1162/neco.2008.12-06-420 PMID: 18085991
8. Schall JD, Palmeri TJ, Logan GD. Models of inhibitory control. Philos Trans R Soc Lond B Biol Sci. 2017;372(1718):20160193. https://doi.org/10.1098/rstb.2016.0193 PMID: 28242727
9. Weigard A, Heathcote A, Matzke D, Huang-Pollock C. Cognitive modeling suggests that attentional failures drive longer stop-signal reaction time estimates in attention deficit/hyperactivity disorder. Clin Psychol Sci. 2019;7(4):856–72. https://doi.org/10.1177/2167702619838466
10. Fosco WD, Kofler MJ, Alderson RM, Tarle SJ, Raiker JS, Sarver DE. Inhibitory Control and Information Processing in ADHD: Comparing the Dual Task and Performance Adjustment Hypotheses. J Abnorm Child Psychol. 2019;47(6):961–74. https://doi.org/10.1007/s10802-018-0504-9 PMID: 30547312
11. Mar K, Townes P, Pechlivanoglou P, Arnold P, Schachar R. Obsessive compulsive disorder and response inhibition: Meta-analysis of the stop-signal task. J Psychopathol Clin Sci. 2022;131(2):152–61. https://doi.org/10.1037/abn0000732 PMID: 34968087
12. White CN, Congdon E, Mumford JA, Karlsgodt KH, Sabb FW, Freimer NB, et al. Decomposing decision components in the stop-signal task: a model-based approach to individual differences in inhibitory control. J Cogn Neurosci. 2014;26(8):1601–14. https://doi.org/10.1162/jocn_a_00567 PMID: 24405185
13. Hanes DP, Schall JD. Countermanding saccades in macaque. Vis Neurosci. 1995;12(5):929–37. https://doi.org/10.1017/s0952523800009482 PMID: 8924416
14. Hanes DP, Patterson WF 2nd, Schall JD. Role of frontal eye fields in countermanding saccades: visual, movement, and fixation activity. J Neurophysiol. 1998;79(2):817–34. https://doi.org/10.1152/jn.1998.79.2.817 PMID: 9463444
15. Casey BJ, Cannonier T, Conley MI, Cohen AO, Barch DM, Heitzeg MM, et al. The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. Dev Cogn Neurosci. 2018;32:43–54. https://doi.org/10.1016/j.dcn.2018.03.001 PMID: 29567376
16. Bissett PG, Logan GD. Selective stopping? Maybe not. J Exp Psychol Gen. 2014;143(1):455–72. https://doi.org/10.1037/a0032122 PMID: 23477668
17. Bissett PG, Hagen MP, Jones HM, Poldrack RA. Design issues and solutions for stop-signal data from the Adolescent Brain Cognitive Development (ABCD) study. Elife. 2021;10:e60185. https://doi.org/10.7554/eLife.60185 PMID: 33661097
18. Kaelbling LP, Littman ML, Cassandra AR. Planning and acting in partially observable stochastic domains. Artific Intell. 1998;101(1–2):99–134. https://doi.org/10.1016/s0004-3702(98)00023-x

19. Rao RPN. Decision making under uncertainty: a neural model based on partially observable markov decision processes. Front Comput Neurosci. 2010;4:146. https://doi.org/10.3389/fncom.2010.00146 PMID: 21152255
20. Yu AJ, Dayan P. Uncertainty, neuromodulation, and attention. Neuron. 2005;46(4):681–92. https://doi.org/10.1016/j.neuron.2005.04.026
21. Shenoy P, Yu AJ, R P R. A rational decision making framework for inhibitory control. Adv Neural Inform Process Syst. 2010;23.
22. Shenoy P, Yu AJ. Rational decision-making in inhibitory control. Front Hum Neurosci. 2011;5:48. https://doi.org/10.3389/fnhum.2011.00048 PMID: 21647306
23. Dayan P, Daw ND. Decision theory, reinforcement learning, and the brain. Cogn Affect Behav Neurosci. 2008;8(4):429–53. https://doi.org/10.3758/CABN.8.4.429 PMID: 19033240
24. Cranmer K, Brehmer J, Louppe G. The frontier of simulation-based inference. Proc Natl Acad Sci U S A. 2020;117(48):30055–62. https://doi.org/10.1073/pnas.1912789117 PMID: 32471948
25. Boelts J, Deistler M, Gloeckler M, Tejero-Cantero A, Lueckmann JM, Moss G, et al. Sbi reloaded: a toolkit for simulation-based inference workflows. J Open Source Softw. 2025;10(4):7754. https://doi.org/10.21105/JOSS.07754
26. Vaswani A, Brain G, Shazeer N, Parmar N, Uszkoreit J, Jones L. Attention is All You Need. Adv Neural Inform Process Syst. 2017;30.
27. Papamakarios G, Murray I. Fast ε-free inference of simulation models with Bayesian conditional density estimation. Adv Neural Inform Process Syst. 2016;29.
28. Auchter AM, Hernandez Mejia M, Heyser CJ, Shilling PD, Jernigan TL, Brown SA, et al. A description of the ABCD organizational structure and communication framework. Dev Cogn Neurosci. 2018;32:8–15. https://doi.org/10.1016/j.dcn.2018.04.003 PMID: 29706313
29. Clark DB, Fisher CB, Bookheimer S, Brown SA, Evans JH, Hopfer C, et al. Biomedical ethics and clinical oversight in multisite observational neuroimaging studies with children and adolescents: The ABCD experience. Dev Cogn Neurosci. 2018;32:143–54. https://doi.org/10.1016/j.dcn.2017.06.005 PMID: 28716389
30. Achenbach T. Manual for the child behavior checklist and revised child behavior profile. University of Vermont; 1983.
31. Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K. Research Domain Criteria (RDoC): Toward a new classification framework for research on mental disorders. Am J Psychiat. 2010;167(7):748–51. https://doi.org/10.1176/appi.ajp.2010.09091379
32. Ma N, Yu AJ. Inseparability of Go and Stop in Inhibitory Control: Go Stimulus Discriminability Affects Stopping Behavior. Front Neurosci. 2016;10:54. https://doi.org/10.3389/fnins.2016.00054 PMID: 27047324
33. Bellman R. On the Theory of Dynamic Programming. Proc Natl Acad Sci U S A. 1952;38(8):716–9. https://doi.org/10.1073/pnas.38.8.716 PMID: 16589166
34. Durkan C, Bekasov A, Murray I, Papamakarios G. Neural spline flows. Adv Neural Inform Process Syst. 2019;32.
35. Beaumont MA, Zhang W, Balding DJ. Approximate Bayesian computation in population genetics. Genetics. 2002;162(4):2025–35. https://doi.org/10.1093/genetics/162.4.2025 PMID: 12524368
36. Rozet F, et al. Zuko: Normalizing flows in PyTorch. 2022. Available from: https://pypi.org/project/zuko
37. Wechsler D. Wechsler intelligence scale for children–fifth edition (WISC-V). Bloomington, MN: Pearson; 2014.
38. Owens MM, Allgaier N, Hahn S, Yuan D, Albaugh M, Adise S, et al. Multimethod investigation of the neurobiological basis of ADHD symptomatology in children aged 9-10: baseline data from the ABCD study. Transl Psychiatry. 2021;11(1):64. https://doi.org/10.1038/s41398-020-01192-8 PMID: 33462190
39. Boucher L, Palmeri TJ, Logan GD, Schall JD. Inhibitory control in mind and brain: an interactive race model of countermanding saccades. Psychol Rev. 2007;114(2):376–97. https://doi.org/10.1037/0033-295X.114.2.376 PMID: 17500631
40. Band GPH, van der Molen MW, Logan GD. Horse-race model simulations of the stop-signal procedure. Acta Psychol (Amst). 2003;112(2):105–42. https://doi.org/10.1016/s0001-6918(02)00079-3 PMID: 12521663
41. Matzke D, Love J, Wiecki TV, Brown SD, Logan GD, Wagenmakers EJ. Release the BEESTS: Bayesian estimation of ex-Gaussian sTop-signal reaction time distributions. Frontiers in Psychology. 2013;4(DEC):69450. https://doi.org/10.3389/FPSYG.2013.00918
42. Matzke D, Dolan CV, Logan GD, Brown SD, Wagenmakers E-J. Bayesian parametric estimation of stop-signal reaction time distributions. J Exp Psychol Gen. 2013;142(4):1047–73. https://doi.org/10.1037/a0030543 PMID: 23163766
43. Weigard A, Matzke D, Tanis C, Heathcote A. A cognitive process modeling framework for the ABCD study stop-signal task. Dev Cogn Neurosci. 2023;59:101191. https://doi.org/10.1016/j.dcn.2022.101191 PMID: 36603413
44. Logan GD, Van Zandt T, Verbruggen F, Wagenmakers E-J. On the ability to inhibit thought and action: general and special theories of an act of control. Psychol Rev. 2014;121(1):66–95. https://doi.org/10.1037/a0035230 PMID: 24490789
45. Tillman G, Van Zandt T, Logan GD. Sequential sampling models without random between-trial variability: the racing diffusion model of speeded decision making. Psychon Bull Rev. 2020;27(5):911–36. https://doi.org/10.3758/s13423-020-01719-6 PMID: 32424622
46. Epstein JN, Karalunas SL, Tamm L, Dudley JA, Lynch JD, Altaye M, et al. Examining reaction time variability on the stop-signal task in the ABCD study. J Int Neuropsychol Soc. 2023;29(5):492–502. https://doi.org/10.1017/S1355617722000431 PMID: 36043323
47. Sebastian A, Forstmann BU, Matzke D. Towards a model-based cognitive neuroscience of stopping - a neuroimaging perspective. Neurosci Biobehav Rev. 2018;90:130–6. https://doi.org/10.1016/j.neubiorev.2018.04.011 PMID: 29660415

48. Brown SD, Heathcote A. The simplest complete model of choice response time: linear ballistic accumulation. Cogn Psychol. 2008;57(3):153–78. https://doi.org/10.1016/j.cogpsych.2007.12.002 PMID: 18243170

49. Verbruggen F, Logan GD. Models of response inhibition in the stop-signal and stop-change paradigms. Neurosci Biobehav Rev. 2009;33(5):647–61. https://doi.org/10.1016/j.neubiorev.2008.08.014 PMID: 18822313

50. Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD. The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. Psychol Rev. 2006;113(4):700–65. https://doi.org/10.1037/0033-295X.113.4.700 PMID: 17014301

51. Gold JI, Shadlen MN. Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. Neuron. 2002;36(2):299–308. https://doi.org/10.1016/s0896-6273(02)00971-6 PMID: 12383783

52. Rakitin BC, Gibbon J, Penney TB, Malapani C, Hinton SC, Meck WH. Scalar expectancy theory and peak-interval timing in humans. J Exp Psychol Anim Behav Process. 1998;24(1):15–33. https://doi.org/10.1037/0097-7403.24.1.15 PMID: 9438963

53. Ma N, Yu AJ. Statistical learning and adaptive decision-making underlie human response time variability in inhibitory control. Front Psychol. 2015;6:1046. https://doi.org/10.3389/fpsyg.2015.01046 PMID: 26321966

54. Ide JS, Shenoy P, Yu AJ, Li CR. Bayesian prediction and evaluation in the anterior cingulate cortex. J Neurosci. 2013;33(5):2039–47. https://doi.org/10.1523/JNEUROSCI.2201-12.2013 PMID: 23365241

55. Huys QJM, Maia TV, Frank MJ. Computational psychiatry as a bridge from neuroscience to clinical applications. Nat Neurosci. 2016;19(3):404–13. https://doi.org/10.1038/nn.4238 PMID: 26906507

56. Montague PR, Dolan RJ, Friston KJ, Dayan P. Computational psychiatry. Trends Cogn Sci. 2012;16(1):72–80. https://doi.org/10.1016/j.tics.2011.11.018 PMID: 22177032