Computational Modeling of Proactive, Reactive, and Attentional Dynamics in Cognitive Control

Percy K. Mistry*[1], Stacie L. Warren[4], Nicholas K. Branigan[1], Weidong Cai[1,2], Vinod Menon*[1,2,3]

[1] Department of Psychiatry & Behavioral Sciences
[2] Wu Tsai Neurosciences Institute
[3] Department of Neurology & Neurological Sciences
Stanford University, Stanford, CA

[4] School of Behavioral and Brain Sciences, Department of Psychology,
University of Texas, Dallas

* Corresponding authors:
Percy K. Mistry, Ph.D. & Vinod Menon, Ph.D.
email: percym@stanford.edu; menon@stanford.edu

Abstract

We developed a novel Proactive Reactive and Attentional Dynamics (PRAD) computational model designed to dissect the latent mechanisms of inhibitory control in human cognition. Leveraging data from over 7,500 participants in the NIH Adolescent Brain Cognitive Development study, we demonstrate that PRAD surpasses traditional models by integrating proactive, reactive, and attentional components of inhibitory control. Employing a hierarchical Bayesian framework, PRAD offers a granular view of the dynamics underpinning action execution and inhibition, provides debiased estimates of stop-signal reaction times, and elucidates individual and temporal variability in cognitive control processes. Our findings reveal significant intra-individual variability, challenging conventional assumptions of random variability across trials. By addressing nonergodicity and systematically accounting for the multi-componential nature of cognitive control, PRAD advances our understanding of the cognitive mechanisms driving individual differences in cognitive control and provides a sophisticated computational framework for dissecting dynamic cognitive processes across diverse populations.

Introduction

Human cognition is a dynamic process, which relies in part, on goal-directed beliefs about task characteristics, context-dependent flexible action control, the capacity to learn from a history of decisions and consequences, leading to moment-to-moment adaptation of response strategies to optimize behavioral outcomes [1-5]. Impairments in cognitive systems that regulate such dynamic processes underlying everyday cognitive functioning are a hallmark of psychopathology [6-10]. Identifying latent cognitive factors that drive adaptive and maladaptive behavioral dynamics is critical for understanding individual differences in how cognitive processes unfold over time, and how their associated alterations in regulatory systems affect symptom presentation in neuropsychiatric disorders [11-18]. However, conventional methods are unable to reveal multi-componential latent constructs that govern dynamic cognitive control processes.

Here, we develop Proactive Reactive and Attentional Dynamics (PRAD), a novel computational model to characterize and measure latent cognitive constructs that govern behavioral dynamics of action execution and inhibition, whose deficits are often associated with multiple psychiatric disorders including attention deficit hyperactivity disorder, autism, substance abuse, and schizophrenia. This model was applied to a response inhibition paradigm in the large-scale (N > 7500) Adolescent Brain Cognitive Development (ABCD)[19] study to uncover the mechanisms by which dynamic cognitive processes involved in response initiation and inhibition are regulated, and the nature of individual differences in the latent cognitive constructs associated with such adaptive and maladaptive regulation.

Inhibitory control, the ability to withhold or cancel undesirable action, thought, and emotion, is fundamental to goal-directed behaviors [20-26]. The stop-signal task (SST, **Figure 1A**) is a widely used paradigm[27,28] to study inhibitory control mechanisms and their neural underpinnings. The SST involves making a response to a *Go* signal but inhibiting the prepared response when the *Go* signal is quickly followed by an infrequent *Stop* signal. The time interval between Go and Stop signals is called the stop-signal delay (SSD) and is experimentally manipulated. On stop signal trials with longer SSD, the prepotent Go response is cognitively further along, and more difficult to stop after detecting the Stop signal. The SST has been used in a variety of domains, including non-human primates [29-31], rodents [32-35], during development in children [20,36-38], through the adult human life span [39,40], in neurodiverse populations [21,22,41-45], psychiatric disorders [23,24,46-48], under the effect of medication [49,50] and intervention[51-53], substance dependence [25,54,55], sleep disorders [56,57], learning difficulties [58,59], eating disorders [26,60], in studies of pregnancy related changes [61], and genetic basis of inhibitory control [62]. Understanding the dynamic cognitive mechanisms that underlie SST holds great promise for enhancing our knowledge of latent processes driving cognitive functioning. Despite its widespread use, the SST and the traditional computational models applied to interpret it face significant challenges that limit their explanatory power and practical utility.

Current computational models of inhibitory control suffer from several limitations. In these models, the efficiency of an individual's inhibitory control is measured by estimating the latent stop-signal reaction time (SSRT). However, the SSRT cannot be measured directly and is typically estimated using a race model [63-65]. One of the primary concerns with traditional SST analysis is the validity of SSRT measures under conditions that violate assumptions of such a

race model, including context and stochastic independence. Recent research[66-68] has highlighted the severe implications of these violations, questioning the reliability of SSRT as a definitive measure of inhibitory control [66,69], although recent approaches have proposed model-based solutions to overcome some of these issues [70,71]. The reliability and validity of inferred SSRT measures become more questionable in the presence of specific behavioral patterns that confound SSRT measurement [69,72,73], potentially leading to systematic bias in conventional measures of SSRT and inhibitory control. This critique underscores a critical gap in the current understanding and measurement of inhibitory mechanisms, pointing to the need for more quantitively precise models that can account for the complex dynamics of cognitive processes.

Existing models of inhibitory control often fail to adequately differentiate between proactive and reactive control mechanisms. Proactive control refers to the anticipation and prevention of impulsive actions through the maintenance of goal-relevant information, whereas reactive control involves the suppression of an action in response to a stop signal [1,2,74].The dichotomy between these processes and their interaction with both top-down and bottom-up regulatory mechanisms remains insufficiently explored in current research paradigms, especially in terms of quantitative measurements of such interacting processes [75,76]. Previous studies have modified the standard SST paradigm to probe the proactive component [77-81], introducing variability in task design, and rendering replication more challenging. While the dynamic belief model has been used to study proactive control in the SST [82,83], it captures trial-wise anticipation of a stop signal but is not feasible for estimating other latent components, such as reactive control [84]. This theoretical gap hinders a comprehensive understanding of the multiple dimensions of inhibitory control and their implications for behavior and cognition.

The concept of nonergodicity further complicates the interpretation of SST data. In nonergodic processes, the statistical patterns observed in a group do not necessarily reflect the patterns of individual cases. Thus, conclusions drawn from group averages can be misleading when applied to individuals. [85-88]. There is growing evidence that psychological phenomena are frequently nonergodic [87]. The acknowledgment of nonergodicity calls for dynamic, individualized approaches to analyzing cognitive data, challenging the conventional reliance on aggregated metrics and static models [89-99].

To address these limitations, we developed an integrated computational model that incorporates dynamic modulation of behavior by multiple latent cognitive processes governing inhibitory control, which allow for complex sequential dependencies, do not make assumptions of context or stochastic independence, and establish quantitative measures for multiple components of reactive and proactive control. The PRAD model was implemented within a hierarchical Bayesian framework, allowing for the estimation of individual-level parameters and trial-level measures. This approach enables the identification and measurement of distinct components that characterize individual differences in inhibitory control, while also accounting for the temporal variability in cognitive processes (**Figure 1B**).

We had five primary objectives in this study. Our first objective was to develop, implement, and validate PRAD, a novel computational cognitive model for inhibitory control that comprehensively accounts for the multi-componential dynamic processes that are not currently represented in extant models (see **SI Table S1** for a complete list of features and **SI text** for

details of known issues that these features tackle as well as comparison of these features to some existing accounts of inhibitory control). PRAD provides a novel comprehensive account of dynamic reactive and proactive inhibition, providing more robust and dissociated measures of individual differences in inhibitory control (**Figure 2**). Our second objective was to demonstrate PRAD's robustness across a wide range of measures and its ability to overcome limitations of conventional race models, including providing debiased estimates of SSRT. Our third objective was to investigate nonergodicity and systemic intra-individual variability in inhibitory control processes. Our fourth objective was to examine the distinct components of inhibitory control, including proactive delayed response (PDR) mechanisms and attentional modulation of stopping (AMS), and their dynamic interactions. Our fifth objective focused on evaluating PRAD's predictive power for performance across a broad spectrum of cognitive domains. This goal sought to validate PRAD's effectiveness beyond its initial context and explore its potential as a more precise tool for understanding the latent substrates of cognitive variability in standardized assessments.

We demonstrate that the PRAD model effectively captures the latent processes of inhibitory control with a three-factor structure: proactive control, reactive control, and attention modulation of reactive control. The model's proactive control component significantly accounts for individual performance differences and adaptivity to errors and changing conditions. PRAD's cognitive control parameters also outperform traditional SSRT measures in explaining individual differences across various cognitive tasks, provide debiased measures of inhibitory control, and provide a better explanation of behavioral dynamics in children with diverse cognitive and demographic profiles. These findings validate the PRAD model and highlight its potential for clinical research and use in probing neural instantiations of cognitive dynamics.

Results

PRAD model overview

The Proactive Reactive and Attentional Dynamics (PRAD) model integrates multiple components of inhibitory control, extending beyond traditional approaches. **Figure 2A** illustrates the model's structure, incorporating proactive, reactive, and attentional mechanisms. The go process is modulated by PDR, which is driven by a cognitive state switching mechanism and belief updating about stop signal delays. The stop process combines baseline reactive inhibition with attentional modulation effects. **Figure 2B** and **SI Table S2** illustrate the mechanisms and summarize the key model parameters. Additional analysis demonstrated that the model shows strong parameter recovery (**SI Figure S1**). We applied PRAD to SST data from 7787 individuals from the ABCD dataset.

PRAD reveals distinct cognitively plausible components of inhibitory control

The model envisages a core reactive inhibitory process, modulated by attentional variations in stopping expectancies, and a dynamically adjusted proactive delayed response of the go process, all three of which affect observed inhibitory behavior (**Figure 3A**). Factor analysis of a subset of model parameters that are theoretically relevant to these three aspects of inhibitory control

revealed the robustness of this three-factor structure (CFI 0.997, TLI 0.977, RMSEA 0.045; **Figure 3B, SI Table S3**). Factor 1, representing proactive control, loaded heavily on parameters $(\theta_0, \theta_1, \mu)$ that governed the trial level probability of engaging PDR mechanisms, and the adaptive belief updating about historical stop-signal delays that modulated the PDR duration. Factor 2, capturing basic reactive control (baseline SSRT), was dominated by parameters (δ_s, α_s) governing the baseline stop process. Factor 3, reflecting attentional modulation that influences variations in trial-level SSRT, loaded strongly on parameters (γ_0, γ_1) influencing the dynamic adjustment of stopping expectancy, which in turn affects dynamic SSRT. Factors 1 and 3 were not correlated, but factor 2 shows low correlations with the other 2 factors (|r| = 0.085, 0.16; both $p < 0.0001$), suggesting related but distinct processes. Supplementary analysis (**SI Table S4**) reports control analysis showing that 1- and 2-factor models were not adequate, and a 4-factor model was not identifiable.

The three factors explained a significant portion of individual differences in inhibitory control performance. Regression analyses using all individual-level model parameters significantly predicted various measures of inhibitory control, including PRAD stop-signal reaction time (SSRT; $R^2 = 86\%$), observed stop-failure rate (SFR; $R^2 = 65\%$), observed mean experienced stop-signal delay (xSSD; $R^2 = 77\%$), SSRT coefficient of variation (SSRTCV; $R^2 = 89\%$), as well as observed RT ($R^2 = 67\%$) and RT CV ($R^2 = 68\%$), all $p < 0.0001$. Importantly, parameters from each of the three factors contributed uniquely and significantly to these predictions, with standardized beta coefficients ranging from -0.86 to 0.80 (see **SI Table S5**).

These results align with the theoretical constructs of proactive control, reactive control, and attentional modulation, and validate the model's ability to dissociate different aspects of inhibitory control.

PRAD inferred SSRT shows systematic compensation for known biases in traditional SSRT

PRAD-inferred SSRT demonstrated systematic compensation for known biases in traditional SSRT estimates[28,67,69]. Comparing model-based median SSRT (mean 340ms, SD 155ms) to non-parametric integration method SSRT (iSSRT, mean 302ms, SD 135ms) revealed that traditional iSSRT estimates were lower by 38ms on average ($t(7786) = -32.3$, $p < 0.0001$), and showed a smaller degree of individual differences (SD lower by 20ms) between individuals ($F(7786,7786) = 0.76$, $p < 0.0001$).

Crucially, PRAD SSRT estimates were significantly higher for conditions known to lead to underestimation of traditional SSRT (**Figure 4A-E; SI Table S6**). These included participants with higher right skew of RT ($F(3,7783) = 21.9$, $p < 0.0001$), larger RT slowdown ($F(3, 7783) = 9.3$, $p < 0.0001$), high stop success rates ($F(2,7784) = 76.8$, $p < 0.0001$), higher go-omission rates ($t(7785) = 3.4$, $p < 0.0001$) and those classified as race-model violators ($t(7785) = 28.9$, $p < 0.0001$).

Additionally, PRAD revealed systematic differences in previously unexamined conditions, such as high RT variability ($F(3,7783) = 9.2$, $p<0.0001$), high RT kurtosis ($F(3,7783) = 25.4$, $p<0.0001$), high SSRT variability ($F(3,7783) = 116$, $p<0.0001$), and low correlations between SFR and SSD ($F(3,7783) = 72.6$, $p<0.0001$) (**Figures 4F-I**).

These results demonstrate PRAD's ability to provide more accurate and unbiased SSRT estimates across a wide range of performance patterns, particularly for individuals who deviate from typical performance profiles.

PRAD tracks aggregate patterns in overt behavioral measures, outperforming conventional methods

The PRAD model is implemented as a generative hierarchical Bayesian model, which generates posterior distributions for trial-level behavior for each individual based on the inferred parameters. PRAD demonstrated robust performance across various measures and outperformed control models (**Figure 5A-F; SI Figure S2)**. **SI Table S7** shows key observed measures at an aggregate level, and the summarized PRAD model posterior values corresponding to these observed measures.

At the group level, PRAD accurately captured trends in stop failure rates and reaction times across different experimental contingencies (SSD, nSSD). Specifically, the posterior predictives generated by the PRAD model explained aggregate behavioral patterns including (i) non-linear S-shaped patterns of stop failure rate with increasing SSD (**Figure 5A**) and non-monotonicity at low SSD values, an indicator of violations of context independence [67]; (ii) linear increase in stop-failure RT with increasing SSD (**SI Figure S2A**), an indicator of the link between slower RT and better stopping performance, or the influence of SSD on RT; (iii) increasing variability in stopping accuracy (**Figure 5B,D,F**) with numbers of trials since encountering the last stop signal (nSSD), (iv) increasing variability in RT with SSD (**SI Figure S2B**), and (v) lower choice accuracies at low SSDs (**SI Figure S2C**).

We also compared PRAD versus two control models – fixed stopping model (FSM) and random variability model (RVM); details of both are provided in the **Methods**. Crucially, the PRAD model outperformed both control models (FSM, RVM) on all of these aggregate measures. Importantly, control models failed to effectively capture the diverse patterns across different subgroups based on simple observed measures, such as whether SFR increased or decreased with SSD (**Figures 5C,E**) and nSSD (**Figures 5D,F**), while the PRAD model provides superior model fits for every single subgroup.

RMSE distance from the observed aggregate curves (**Figure 5G, SI Table S8**) show that across trends based on different experimental contingencies, PRAD reduced RMSE by between 77% - 80% for SFR, 13% - 45% for RT on stop failures, and 48% - 61% for choice accuracies on stop failure, compared to the FSM model. Similarly, PRAD reduced RMSE by between 58% - 66% for SFR, 45% - 50% for RT on stop failures, and 48% - 51% for choice accuracies on stop failure, compared to the RVM model.

These results suggest that incorporating sequential adaptive processes is crucial for characterizing behavioral dynamics on the SST, which PRAD achieves better than conventional approaches.

PRAD captures behavioral dynamics at the individual-subject level

We then examined PRAD robustness at the individual participant level. We assessed how well the model posterior values generated by PRAD explain behavioral patterns at an individual level (mean values per individual). Individual subject-level fits showed strong correlations between PRAD model fits and observed data for multiple behavioral measures (**SI Figures S3**). Notably, individual-level comparisons showed stronger correlations (**Figure 5H; SI Table S9)** and lower RMSE (**SI Table S10)** between observed and predicted values for PRAD across almost all key measures compared to the RVM and FSM control models. This included go reaction time ($r = 0.91$; $p < 0.0001$), stop-failure reaction time ($r = 0.90$; $p < 0.0001$), and stop-failure rate ($r = 0.85$; $p < 0.0001$), but importantly, also second order effects like post-go RTs ($r = 0.92$, $p < 0.0001$), post-stop RTs ($r = 0.86$, $p < 0.0001$), difference between stop and go RTs ($r = 0.51$, $p < 0.0001$), post-inhibitory differences in RTs ($r = 0.45$, $p < 0.0001$), and dynamic within-subject correlations like SFR vs SSD ($r = 0.63$, $p < 0.0001$) and SFR vs nSSD ($r = 0.73$, $p < 0.0001$).

For a significant proportion of participants, observed and latent measures like RTs and SSRT show within-subject correlations with various sequential or experimental contingencies **(SI Table S11)**. Analysis of intra-individual variability in Go RTs and SSRTs revealed that significant variance could be explained by model parameters, rather than random noise (**SI Figures S4**). For reaction time coefficient of variation (RTCV), 72% of the variability was explained by model parameters. For SSRT coefficient of variation (SSRTCV), 81% was explained by model parameters. PRAD's ability to account for the wide range of variability and individual differences is reflected in the Kullback-Leibler divergence between the distribution of observed RT related measures and model posterior values of these observed measures. PRAD reduces the KL divergence (thus providing a closer match to the observed range of individual differences) for RT related distributions by between 29%-89% compared to RVM and 4%-83% compared to FSM (**SI Table S12**). These results demonstrate that PRAD can accurately capture behavioral variability at both the individual-subject level and in terms of individual differences between subjects. Supplementary analysis shows that a large proportion of this explained variance is attributable to novel adaptive model parameters. See **SI text** for additional details on intra-individual variability and estimates of PRAD latent dynamic measures (**SI Table S13).**

Additionally, we measured Deviance information criteria (DIC), which assesses model fit appropriately penalized for model complexity. In spite of the additional complexity, PRAD resulted in lowest (best) DIC values for 60% of the individuals, compared to 22% for the RVM, and 18% for the FSM model, suggesting that the additional complexity was necessary to explain behavior in a majority of individuals.

Visualizing trial-by-trial variability in PRAD components

To visualize the dynamic interplay of key PRAD components and their trial-by-trial variability at the individual subject level, we examined detailed time courses of model-derived measures for representative participants. **Figure 6** illustrates the dynamics involved in stop trials from a single participant, showing how stopping expectancy, SSRT, PDR, and RT interact on a trial-by-trial basis. Stopping expectancy, modulated by attentional regulation, fluctuated considerably from trial to trial and showed an inverse relationship with SSRT. This revealed how dynamic attentional processes can influence inhibitory performance on a moment-to-moment basis. We

also observed that task difficulty (represented by SSD) and stopping efficiency (SSRT) varied substantially across trials, with their sum (SSD+SSRT) providing insight into the overall challenge of inhibition on each trial. Our analysis of stop-failure trials revealed complex relationships between PDR, SSD+SSRT, and RT, demonstrating how proactive and reactive mechanisms interact to determine inhibitory outcomes.

To further elucidate these dynamics, we closely examined three specific stop trials with varying levels of difficulty. We found that successful inhibition could occur even on more difficult trials (higher SSD) when compensatory mechanisms like increased stopping expectancy or heightened PDR were engaged. Conversely, easier trials could result in failures when these compensatory mechanisms were absent.

These observations highlight PRAD's capacity to capture and explain the substantial intra-individual variability in inhibitory control processes, accounting for the complex interplay between task parameters, attentional modulation, and proactive control strategies that occur on a trial-by-trial basis.

PRAD reveals nonergodicity in behavioral dynamics

PRAD revealed substantial evidence for nonergodic processes in inhibitory control. Specifically, the model uncovered differences between within-subject and between-subject correlations for both latent and observed measures **(Figures 7A-D**), indicating nonergodicity. For example **(SI Table S14**), we observed opposite patterns of within-subject and between subjects correlations, in the associations between mean experienced stop-signal delay and stop-failure rate (average within r = 0.22, between = -0.77), RT and probability of proactive state (average within r = 0.42, between = -0.05), RT and SSRT (average within r = 0.56, between = -0.13), stop-failure rate and probability of proactive state (average within r = -0.31, between = 0.06), SSRT and average proactive delay (average within r = 0.01, between = -0.39).These divergences suggests that within-subject and between-subjects effects can differ significantly, reflecting nonergodicity of behavioral dynamics.

Proactive Delayed Response (PDR) mechanisms in inhibitory control

PRAD revealed the substantial contribution of PDR mechanisms to inhibitory control. Across individuals, proactive cognitive states occurred on an average of 78% of trials (2.5th to 97.5th percentile 3-98%), with PDR accounting for approximately 28% (2.5th to 97.5th percentile 2-51%) of average reaction times of 515ms (mean PDR = 150ms, 2.5th to 97.5th percentile 8-322ms; **Figure 8A**).

Between subjects, the average PDR is associated with both the mean RT (r = 0.59, p < 0.0001) of responses as well as RTCV across trials (r = -0.51, p <0.0001), and is negatively correlated with SFR (r = -0.59, p < 0.0001), and positively correlated with xSSD (r = 0.61, p <0.0001). The variability in PDR (PDRCV) is correlated with RTCV (r = 0.36, p <0.0001). Linear regressions (**SI Table S15**) show that across individuals, PDR is significantly related to SFR ($R^2 = 0.48$, $\beta = -0.44$, $p < 0.0001$) and xSSD ($R^2 = 0.46$, $\beta = 0.48$, $p < 0.0001$) even after controlling for SSRT. The results demonstrate that individuals with longer delayed responding show more

successful inhibitory control (lower SFR, higher xSSD) by appropriate modulation of RTs, even after controlling for the effect of SSRT on successful inhibition. Regression controlling for the influence of average experienced SSD and SSRT shows that the average probability of proactive states is significantly related to SFR ($R^2 = 0.65,\ \beta = -0.15,\ p < 0.0001$). Within-subject regressions also show that after controlling for SSRT, SSD, and variable drift rate across trials, the probability of proactivity, but not the length of proactive delay, is significantly related to stop-failures (**SI Figure S5**).

Individual differences in PDR were characterized by three key parameters: baseline proclivity for proactive control (θ_0), degree of adaptive or maladaptive error monitoring (θ_1), and persistence of belief updating that affected the tracking of stop-signal delays (μ).

Higher baseline proactivity (θ_0) was positively associated with increased PDR (r = 0.44, p < 0.0001) but not so with RT (r = -0.07, p<0.0001), suggesting that PDR mechanisms may be linked to core capabilities, and individuals with faster processing speeds may also have improved top-down PDR regulation. Baseline proclivity for proactive control (θ_0) is also correlated to lower RT coefficient of variability (r = -0.29, p <0.0001).

Persistence in belief updating (μ) affected the absolute error between tracked and current stop-signal delays (r = 0.42, p < 0.0001), with lower μ values associated with higher recency bias, and more accurate dynamic SSD tracking, and hence more well-calibrated PDR (**Figure 8B**).

Error monitoring (θ_1) influenced the adaptivity of PDR following errors (**Figures 8C-D**), with adaptive individuals (θ_1 < 0) showing an increase in probability of PDR following stop-failures, while maladaptive individuals (θ_1 > 0) showed a decrease. For individuals with $\theta_1 < 0$ (62% of individuals), the delayed response states post stop-failure increases (adaptively) to 90% compared to 73% post go-omission. For individuals with $\theta_1 > 0$, delayed response states post stop-failure are (maladaptively) 55%, compared to 77% post go-omission. The resulting differences in delayed response states manifest as a difference in reaction times (since delayed responding increases RT), between post stop-failure and post go-omission trials. This difference depends on whether individuals demonstrate adaptive (mean increase of 37ms) or maladaptive (mean decrease of 25ms) values of θ_1. Higher values of θ_1 lead to higher go-omission rates (r = 0.527, p < 0.0001), and lower values of θ_1 lead to higher post-inhibitory error effects, with a negative correlation between θ_1 and the difference between post-stop error and post-go error RTs (r = -0.38, p < 0.0001).

These results demonstrate that individual differences in average PDR levels as well as PDR variability are influenced by differences in baseline cognitive states, persistence in belief updating, and error monitoring processes. These results also showcase how the model captures post-inhibitory effects[100] by adaptive or maladaptive modulation of the proactive inhibitory channels.

Attentional Modulation of Stopping (AMS) mechanisms in inhibitory control

Attentional Modulation of Stopping (AMS) emerged as a crucial component of inhibitory control in the PRAD model. AMS is governed by two key parameters: attention-based adaptivity (γ_1)

and attention length ($\gamma 0$). These parameters modulate the stopping expectancy which in turn affects the SSRT on a trial-by-trial basis.

Across participants, the average stopping expectancy ranged from 0.43 to 0.66 (2.5th to 97.5th percentile). Across all trials and participants, the stopping expectancy ranged from 0.38 to 0.71 (2.5th to 97.5th percentile). This stopping expectancy is the implicit expectation about the probability of encountering a stop signal; it is the initial bias of the stopping drift diffusion process, modulating the baseline reactive stopping process. Linear regression **(SI Table S16**) reveals that after controlling for stop process drift, response thresholds, and non-decision time (i.e., the remaining parameters affecting SSRT), the AMS driven stopping expectancy still has an influence on SSRT ($R^2 = 0.84,\ \beta = -0.21,\ p < 0.0001$).

The variability in stopping expectancy represents the strength of the AMS effect, and the coefficient of variation of stopping expectancy ranged from 0.005 (2.5th percentile; weak AMS effect) to 0.229 (97.5th percentile; moderate AMS effect). The SSRTCV is positively correlated with variability in stopping expectancy (r = 0.55, p <0.0001). Linear regression (**SI Table S16,** $R^2 = 0.83$) *s*hows that average stopping expectancy ($\beta = 0.36,\ p < 0.0001$) and coefficient of variation of stopping expectancy ($\beta = 0.55,\ p < 0.0001$) have a significant influence on SSRTCV even after controlling for the effects of stop process drift rate and stop process decision threshold.

To further quantify the impact of attentional modulation on stop process dynamics across our sample, we analyzed subgroups based on their attentional modulation parameter (γ_1). Analysis revealed that 57% of participants showed decreasing attentional control ($\gamma_1 < 0$) as the number of trials since the last stop signal increased, while 43% showed increasing control ($\gamma_1 > 0$). Here increasing attentional control refers to increasing bias of the stopping process, which on the presentation of a stop signal stimulus would result in faster SSRTs.

We found significant differences between these subgroups in several key measures as a function of the number of trials since the last stop signal (nSSD). Changes in stopping expectancy with nSSD (**Figure 9A**) showed a divergent trend between subgroups (t(7785) = 839, p < 0.0001), with the $\gamma_1 > 0$ group maintaining higher expectancy as nSSD increased (average r = 0.96), while the $\gamma_1 < 0$ group showed decreasing expectancy (average r = -0.97). This translated to significant differences in the correlation of SSRT **(Figure 9B)** and nSSD between subgroups (t(7785) = -157, p < 0.0001), and correlation of observed SFR **(Figure 9C)** and nSSD between subgroups (t(7785) = -52, p < 0.0001), with the $\gamma_1 > 0$ group showing more stable performance, in terms of reducing SFR with nSSD (average correlation -0.21 vs 0.16). Thus, AMS parameters significantly influenced patterns of stop-failure rates.

AMS processes also provided a continuous process explanation for what have previously been termed trigger failures [101-103], or assumed failures to initiate the stopping process **(Figure 9D)**. The AMS mechanism in the PRAD model allowed the distribution of SSRTs to include higher valued SSRTs on some trials. As a result, the total SSD + SSRT were higher than the 90th percentile individual RT for 14% of stop trials across individuals (95% CI from 0% to 82%; average 5% for stronger factor 2 and 22% for weaker factor 2; **SI text** for more details). The AMS modulation, which allows for greater trial-level variability in the SSRT thus results in a

higher degree of SSRT being higher than the typical RTs, especially for participants low on factor 2 (weak reactive inhibition), as seen in **Figure 9D**. Such trials, in the absence of the PRAD model would be difficult to explain and be classified as trigger failures within the traditional account.

The distribution of SSRT and RT (**Figure 9E**) across the sample revealed substantial individual differences, with SSRT showing greater between-subject variability (CV = 0.44) compared to RT (average CV = 0.16). Finally, we found that the log ratio of SSRT/RT (**Figure 9F, SI Table S16**) was significantly influenced by both stop process drift rate ($\beta = -0.27$, $p < 0.0001$) and threshold ($\beta = 0.63$, $p < 0.0001$), accounting for 74% of the variance in this measure.

These findings demonstrate the pervasive influence of attentional modulation on stop process dynamics across our sample, highlighting its role in explaining individual differences in inhibitory control performance. See **SI text** for additional details on AMS mechanisms and trigger failure explanations.

Dynamic interactions between proactive and reactive control processes

PRAD uncovered complex interaction dynamics between proactive and reactive control mechanisms. **Figure 10** demonstrates the variability in PDR and AMS processes across individuals, highlighting how different combinations of these processes can lead to similar overall performance but through distinct cognitive mechanisms. **SI text** and **SI Figure S6** further elucidate these interactions between proactive and reactive processes.

These results highlight the importance of considering the dynamic interplay between different control mechanisms when characterizing individual differences in inhibitory control.

PRAD better predicts cognitive performance on multiple NIH Toolbox tasks

PRAD parameters demonstrated superior predictive power for performance on NIH Toolbox cognitive tasks compared to traditional SSRT measures. Using support vector machines to fit the overall NIH cognitive toolbox scores, we found that adjusted R^2 improved from 6.3% to 27.7% when using PRAD parameters instead of traditional iSSRT measures (**Figure 11A)**. The correlations between actual and fitted values increased from 0.25 to 0.53 (**Figure 11B)**. Cross-validated results confirmed this pattern, with PRAD parameters maintaining higher predictive power (adjusted $R^2 = 15.3\%$; correlations 0.40) compared to iSSRT (adjusted $R^2 = 4.6\%$; correlations 0.22) (**Figure 11C-D)**.

Similar improvements were observed for individual NIH cognitive toolbox tasks and subscores, including Flanker Inhibitory Control, Dimensional Change Card Sort, and Pattern Comparison Processing Speed tasks (**SI Table S17; Figure 11**).

These findings suggest that PRAD parameters capture generalizable aspects of cognitive control that extend beyond the specific context of the stop-signal task, providing a more comprehensive characterization of individual differences in cognitive abilities.

Reducing biases in inferences about nontypical subpopulations

PRAD provides improved fits to data compared to the control models. For some of the key measures – RT, RT SD, SFR, and correlation between SFR and SSD – we evaluated the absolute value of residuals based on both the PRAD and RVM models (**SI Table S18**). The effect sizes (Cohen's d) for improved fits on these measures for PRAD vs RVM range from 0.33 to 1.27.

For each measure, we then evaluated the mean absolute residuals by population subgroups based on cognitive ability (median split using NIH cognitive toolbox scores), age (median split), and family income (based on income lower than or greater than $50k annually). We computed the bias against nontypical subgroups by comparing these mean absolute residuals between the nontypical and typical subgroups (lower vs higher cognitive ability, younger vs older children, lower vs higher family income).

While both PRAD and RVM models demonstrated bias (relatively better fits for typical vs nontypical subgroups), this bias was significantly lower in the PRAD model vs RVM for RT, RT SD, and correlation between SFR and SSD. The percentage reduction in bias (**SI Figure S7**) across these three measures ranged from 27%-65% (Cohen's d 0.16 – 0.25) for cognitive subgroups, 39%-68% (Cohen's d 0.07 – 0.23) for age based subgroups, and 24% to 71% (Cohen's d 0.09 – 0.21) for income based subgroups. Importantly, to identify a reduction in bias, we ensured that the PRAD model showed lower mean absolute value of residuals for both typical and nontypical subgroups compared to RVM and also showed a reduction in difference between these.

Discussion

We developed and validated a novel computational model of Proactive, Reactive, and Attentional Dynamics (PRAD), that characterizes latent proactive, reactive, and attentional components underlying inhibitory control. We leveraged a very large dataset (N > 7,500) of children ages 9-10 from the NIH ABCD study, which allowed us to probe distinct sources of response intraindividual variability in ways that were previously not possible. The ABCD study provides a unique opportunity to investigate cognitive processes at an unprecedented scale, with a sample size that far exceeds most previous studies in the field. This large dataset enabled us to conduct detailed analyses at both the group and individual levels, as well as to examine cognitive dynamics at the single-trial level. PRAD demonstrates robustness to violations of context independence, a limitation of extant models. PRAD model parameters provide a better explanation of individual differences in performance across a range of executive function, attention, processing speed, language, and learning/memory tasks compared to conventional models. We delineate specific mechanisms of proactive control and attention modulation, demonstrating their interaction and ability to compensate for weak reactive inhibitory control. PRAD as a sophisticated, multicomponent, model offers a dynamic framework for precisely characterizing goal-directed behaviors and meaningfully delineating individual differences in cognitive control processes and their functional consequences. These advances will be critical for examining dynamic neurocognitive mechanisms of inhibitory control in diverse populations.

PRAD model provides a robust framework for understanding inhibitory control dynamics

Our first goal was to develop a new computational model of response inhibition which explicitly accounts for proactive control, overcomes limitations of conventional race models of reactive control, and explicitly measures sources of intraindividual response variability rather than treating it as random noise. We demonstrate that PRAD provides a strong fit to the data at both the aggregate and individual levels. Linear regressions showed that nearly all model parameters (**SI Table S5**) significantly predicted conventional measures of inhibitory control and sources of variance known to affect it. Our model involved three components of inhibitory control: a basic Go and reactive control process, a dynamic proactive delayed response mechanism, and a dynamic attention modulation mechanism (**Figure 2**). Validating the model's theoretical constructs, a factor analysis revealed three interpretable factors - proactive control, reactive control, and attention modulation (**Figure 3**). These results demonstrate that PRAD is a robust cognitive framework for distinguishing among three dissociable pathways of response inhibition beyond reactive control alone [104-106]. PRAD parameters that comprise each latent pathway reliably capture individual differences to more precisely delineate the emergent function of cognitive control and contribute towards the development of a richer theoretical framework [107].

PRAD overcomes limitations of previous approaches and debiases SSRT estimates and inferences about non-typical populations

Traditional SSRT estimates are prone to biases when assumptions of context and stochastic independence are violated. PRAD does not rely on these assumptions, allowing for sequential processes to modulate both Go and Stop processes. Apart from the fact that traditional measures of SSRT cannot capture trial-level variability in SSRT or related cognitive dynamics, another big limitation is that they have been shown to be biased and unreliable under certain conditions. Specifically, it has been shown that estimates based on the independent race model can be biased [67], and suffer from underestimation – particularly when participants show a higher right skew of RT[28,69], when participants show a larger slowdown in RT [69], when stop success rates are very high, when go omission rates are higher [28], and when stop RTs are longer than go RTs [28]. The model's robustness was tested against conventional SST estimates using the heterogeneous ABCD dataset, which included performance patterns that violate race model assumptions. PRAD effectively compensated for known, systematic biases in conventional measures, particularly when stop success rates exceeded 75%, go-omission rates exceeded 20% and for violators of context independence.

Specifically, PRAD revealed that the degree of SSRT underestimation in race models was larger for participants with slower reaction times, greater RT slowdown, higher stop success rates, higher go-omission rates, and for violators of context independence. Results demonstrated that PRAD robustly identifies and characterizes response inhibition even when assumptions of the race model and context independence are violated, effectively compensating for known, systematic biases in conventional measures. Thus, PRAD effectively characterizes and rectifies biases inherent in traditional cognitive control models. This is especially important when investigating neurodevelopmental disorders and older adults who exhibit diminished performance and greater variability.

The detection of other systematic patterns of differences from non-parametric methods (**Figure 4**), while exploratory, is indicative that there may be other previously unexamined specific patterns of behavior that lead to biases in traditional non-parametric iSSRT estimates, that can be detected by the PRAD models.

Additionally, compared to control models, PRAD more effectively captures the full spectrum of cognitive heterogeneity present in both typical and non-typical or under-represented developing populations, as demonstrated by the reduction in bias of inferences (**SI Figure S7**) for nontypical subgroups based on cognitive abilities, age, and family income.

Nonergodicity in cognitive dynamics revealed by PRAD

Nonergodicity in a behavioral context occurs when statistics of a behavior over time (within-individual dynamics) do not converge to statistics of the behavior over individuals (between-individual dynamics) [108]. In other words, nonergodic processes exhibit different inferences when behavioral dynamics are analyzed at the within-individual level compared to the between-individual level. This distinction is crucial because within-subjects conclusions are often drawn from between-subjects inferences [109]. Yet, such generalizations are only valid for ergodic processes [110].

In our study using the SST and the PRAD model, we found strong evidence for nonergodicity in cognitive dynamics related to response inhibition. Specifically, we observed that the relationships between different cognitive processes, such as proactive control, reactive control, and attentional modulation, exhibited different patterns when analyzed within individuals over time compared to between individuals (**Figure 7**). Moreover, the relationship between probability of proactivity and stop-failure rate showed opposite patterns within and between individuals (**Figure 7**). These findings suggest that the interactions among cognitive processes underlying response inhibition are nonergodic, and that within-individual dynamics cannot be fully captured by between-individual analyses.

The implications of these findings are twofold. First, they highlight the importance of considering individual differences in cognitive dynamics when studying response inhibition and other cognitive processes. Group-level analyses may not adequately capture the complex, time-dependent relationships between cognitive processes within individuals. Second, the presence of nonergodicity along with individual differences in within-subject dynamics suggest that personalized approaches to understanding and modifying cognitive control deficits may be necessary. Interventions targeting specific cognitive processes, such as proactive control or attentional modulation, may have different effects depending on an individual's unique cognitive dynamics.

Integrated modeling of proactive inhibitory control

Proactive control, the ability to anticipate and prepare for forthcoming events, is often neglected in conventional models. PRAD explicitly models proactive control using dynamic parameters that capture trait-like and state mechanisms distinguishing proactive delayed responses on a trial-by-trial basis. A key finding is the substantial contribution of proactive control processes, beyond

just reactive control, in shaping inhibitory performance. PRAD revealed that across individuals, over 75% of trials engaged some form of proactive control, with an average delay of 150ms constituting nearly 30% of average reaction times. Crucially, longer proactive delays were associated with more successful inhibition, underscoring the pivotal yet underappreciated role of proactive control in SST performance. This echoes a few previous findings that revealed negative association between proactive control and SSRT and suggested that better proactive control is related to faster stopping speed [79,81]. It is noteworthy that these previous studies relied on additional experimental manipulation to probe proactive control whereas PRAD can identify proactive control components from the standard SST. Such studies with experimentally manipulated versions of the SST task have estimated go response delays in the range of 100-140ms[79,111,112], similar to the PRAD inferred average of 150ms

PRAD's latent parameters shed light on how individual differences in baseline proactive control tendency, tracking of stimuli in memory, and adaptive vs maladaptive error monitoring dynamically influence the manifestation of proactive control. Baseline proactive control tendency was the largest contributor to overall proactive delayed responding, followed by working memory and error sensitivity. Memory recency bias correlated with individual variability of proactive delay. PRAD differentiated adaptive and maladaptive error response patterns, which manifested as differences in reaction times. Such patterns likely have utility in characterizing cognitive control subtypes in typical development and psychopathology.

These findings have important implications. Previous models largely focused on reactive control, which is important for responding to unexpected stimuli in the environment. However, many daily life situations require proactive control to minimize the need for reactive inhibition and reduce impulsivity. As proactive control is generated by an individual's goals, explicit measurements of it may have greater ecological validity in characterizing everyday response tendencies and psychopathology.

Attentional modulation plays a key role in shaping inhibitory control dynamics

PRAD also revealed that attentional modulation, based on individual differences in latent measures of sustained attention, affects intra-individual variability in SSRT and stop failure rate. We postulated that sustained attention modulates expectancy of stopping, which is influenced by the duration since the last stop signal was encountered. In PRAD, this stopping expectancy is governed by the parameter $\gamma 1$. Negative values of $\gamma 1$ indicate a decline in expectancy and a corresponding decrease in attentional control over the reactive stop process. Conversely, positive values of $\gamma 1$ indicate a rise in expectancy and an amplification of attentional control over the reactive stop process. We found that $\gamma 1$ influences the bias of the stop drift process, thereby modulating the baseline reactive inhibition.

Notably, linear regression revealed that even after controlling for stop process drift, response thresholds, and non-decision time, attentional modulation, as indexed by γ_1, still has a significant influence on SSRT. The average attentional modulation was found to significantly affect variability in SSRT and was correlated with SSRT coefficient of variation. Furthermore, there was a notable within-individual correlation between stop-failure rates and trials since the last stop trial, highlighting the dynamic nature of attentional modulation on inhibitory control.

These findings underscore the pivotal role of attentional modulation and individual differences in latent measures of sustained attention in shaping intra-individual variability in SSRT and stop failures. Our results advance the understanding of how the dynamics of inhibitory control are influenced by stopping biases and highlight the intricate interplay between attentional modulation and inhibitory control. The ability of the PRAD model to capture these complex relationships provides a more comprehensive and nuanced view of the cognitive processes underlying response inhibition, which may have important implications for understanding inhibitory control deficits in various clinical populations.

PRAD better predicts cognitive task performance on a wide range of NIH toolbox tasks

The PRAD model significantly enhances our ability to predict cognitive task performance across a spectrum of domains assessed by the NIH Toolbox Cognitive Battery. Specifically, our analysis revealed that parameters derived from the PRAD model outperformed traditional SSRT metrics for explaining individual differences in tasks measuring executive function, attention, processing speed, language, and learning/memory (**Figure 11**). The NIH Toolbox is an integral component within the Research Domain Criteria (RDoC) framework, utilized for assessing a wide range of cognitive functions across various disorders and developmental stages. This comprehensive toolset enables researchers to bridge cognitive performance with underlying neural and psychological mechanisms, making the PRAD model's predictive power particularly valuable.

Our findings lend substantial external validity to the PRAD model, underscoring its utility in representing latent processes crucial for a broad range of cognitive tasks. This is crucial as the RDoC framework aims to understand psychiatric disorders through a dimensional approach that transcends traditional diagnostic categories. By aligning PRAD with tasks from the NIH Toolbox, our findings highlight the model's capacity to capture cognitive processes that are foundational across multiple domains of function and dysfunction.

Standard neuropsychological assessments often fall short in isolating specific control processes that contribute to cognitive task performance. This limitation is addressed by the PRAD model's process specificity, which enables a more nuanced characterization of control deficits. For instance, traditional tests might not differentiate between an individual's inherent proactive control capabilities and their reactive control responses under pressure. In contrast, PRAD's detailed parameterization allows for the disentanglement of these processes, offering insights into how specific aspects of control contribute to overall task performance.

Moreover, the PRAD model's ability to predict performance across diverse cognitive domains validated by the NIH Toolbox not only reinforces the model's external validity but also emphasizes its potential in identifying fundamental control mechanisms that are broadly applicable across various cognitive tasks. This advance sets the stage for future research aimed at integrating cognitive modeling with clinical diagnostics and therapeutic interventions, guided by the principles of the RDoC framework.

Implications for probing heterogeneity and intraindividual variability

Our modeling and insights have important implications. The availability of dynamic trial-level latent cognitive parameters allows for rigorous quantitative investigations of nonergodic neural processes involved in inhibitory control. The key factors tackled in our work – debiased SSRT measurements for extreme performers, debiased inferences for non-typical or under-represented subgroups, heterogeneity in intraindividual variability, and heterogeneity in proactive, top-down, and bottom-up regulatory processes contributing to modulation of inhibitory control may become even more important in clinical and neurodiverse populations.

Moreover, the distinction between reliance on proactive versus reactive control mechanisms, as illuminated by the PRAD model, offers a refined lens through which maladaptive behaviors and transdiagnostic symptoms can be understood. Individual differences in these control strategies could account for the wide variability in cognitive performance and behavioral outcomes observed across and within psychiatric disorders. This understanding holds significant promise for tailoring interventions to target specific cognitive control deficits, moving towards a more personalized approach in clinical practice. For example, an overreliance on reactive control and a diminished capacity for proactive control may contribute to impulsive behaviors and difficulty with goal-directed planning, which are common features of many psychiatric disorders. By considering individual differences in these control mechanisms, PRAD can help to elucidate the cognitive processes that may underlie common symptoms across different diagnostic categories, in line with the RDoC approach. More broadly, PRAD exemplifies the utility of computational approaches in unraveling the complex dynamics of cognition. It paves the way for developing models that can better capture heterogeneity in cognitive processes across populations, advancing our understanding of mechanisms underlying adaptive and maladaptive behavior.

Conclusion

PRAD provides a powerful computational framework for dissecting the intricate dynamics of inhibitory control in human cognition. The PRAD model effectively characterizes latent proactive, reactive, and attentional components underlying inhibitory control, offering significant implications for understanding the interplay between proactive and reactive control mechanisms. The strong influence of proactive control processes on behavioral variability suggests that factors previously attributed to reactive inhibitory control failures may reflect breakdowns in proactive control. By integrating proactive, reactive, and attentional mechanisms, PRAD advances our understanding of the cognitive underpinnings of inhibitory control and individual differences therein.

The comprehensive and dynamic nature of PRAD provides a robust framework for characterizing cognitive control variations across diverse populations. By leveraging computational models like PRAD, we anticipate advancing our understanding of the mechanistic accounts of cognitive disruptions associated with psychopathology. The availability of dynamic trial-level latent cognitive parameters allows for holistic investigation of nonergodic neural processes involved in inhibitory control, supporting the RDoC framework.

The model's ability to tackle nonergodic processes and systematic biases for non-typical populations enables examining neural mechanisms of control in diverse and clinical groups. PRAD's process specificity and explanatory power highlight its potential for elucidating control deficits in psychopathology and informing individualized interventions. More broadly, PRAD exemplifies the utility of sophisticated computational approaches in unraveling the complex dynamics of cognition. It paves the way for developing models that can better capture heterogeneity in cognitive processes across populations and advancing our understanding of mechanisms underlying adaptive and maladaptive behavior.

Materials and Methods

Participants and inclusion criteria

The present study used a sample of N = 7787 from the Adolescent Brain Cognitive Development (ABCD) study. The ABCD study is an ongoing, longitudinal study within the United States that follows a nationally representative sample of children aged 9–10 at baseline [113,114]. Data were from the baseline visit of the ABCD study[115] (Collection #2573). 9355 participants had two complete runs of the SST task and had data that were successfully fit with the PRAD model, including checks for convergence. We then excluded participants who had a glitch reported in the task presentation (N = 257) and those who had left vs right choice accuracies less than chance level (N = 164). Then, we excluded siblings by randomly keeping one member from each family (using the genetic_paired_subjectid variables from gen_y_pihat; $N = 1147$ excluded). Applying these inclusion criteria left us with a sample of $N = 7787$. See **SI Table S19** for participant demographics.

Stop-Signal Task

Participants completed the SST [116] task during fMRI acquisition. Left and right facing arrows were presented serially as "go" stimuli. Participants indicated the direction of the arrows using a button box and were instructed to respond as quickly and accurately as possible, but to withhold their response on a small subset of trials ("stop" trials) when they saw an upward facing arrow (the "stop" signal), which only appeared after a brief delay (stop signal delay – SSD). Participants completed two runs of 180 trials each, with each run including 30 "stop" trials and 150 "go" trials. On stop trials, the time delay between the "go" and "stop" signals (SSD) was dynamically adjusted by 50 milliseconds increments – increasing after successful stopping and decreasing after unsuccessful stopping, targeted to modulate the difficulty levels so that each participant would be able to successfully inhibit their responses approximately 50% of the time.

Computational Modeling - PRAD

The PRAD model incorporates latent dynamics that respond to endogenous and exogenous variables, with trait measures governing the interaction of such endogenous and exogenous variables with latent processes, giving rise to non-stationary dynamics. This allows the PRAD model to account for violations of context and stochastic independence. Overall, the PRAD model incorporates separate evidence accumulation (drift-diffusion) processes for the go and stop processes, similar to a canonical horse-race model[117]. However, in addition to typical drift-diffusion process parameters, PRAD includes additional individual trait-like and dynamic trial-level measures. The full PRAD model is specified below.

(a) The go process is modeled as a drift diffusion process, with a trial-invariant non-decision time (τ_G) and initial directional bias (β_G), but a trial-varying decision threshold ($\alpha_{G,t}$) and drift rate ($\delta_{G,t}$). The dynamic decision threshold

$$\alpha_{G,t} = \alpha_{G1}\left(1 - \epsilon_{C,t-1}\right) + \alpha_{G2}\left(\epsilon_{C,t-1}\right),$$

where (α_{G1}, α_{G2}) are distinct threshold levels and $\epsilon_{C,t-1}$ is an indicator of whether the left vs right choice on the previous trial was erroneous. Thus, the dynamic threshold implements a form of performance monitoring and varies between two levels based on the outcome of the previous trials, with α_{G1} being the default threshold and α_{G2} reflecting the threshold after post-error adjustments. The dynamic drift rate

$$\delta_{G,t} = \frac{\delta_0 \ \mathbb{S}_{LR,t}}{(1 + \ e^{-12(stim_t - \kappa_0)})} \text{ if } stim_t \neq 0, \text{else } 0.$$

Here δ_0 is a measure of the maximum drift rate for an individual, with the actual drift rate depending on the duration of the go stimulus ($stim_t$) and an individual parameter κ_0, which can be interpreted as the stimulus duration at which the drift rate is half the maximum. $\mathbb{S}_{LR,t}$ assumes values 1 or -1 depending on the direction of the go stimulus (left or right). PRAD allows for trial level changes to the drift rate, overcoming the issues with variable go stimulus durations highlighted in previous work[67].

In addition, in the PRAD model, the onset of the go process may be deliberately delayed in anticipation of a stop signal. This dynamic adaptation is modeled by adding a further delay ω_t to the go process to reflect proactive delayed responding to the go stimulus, where:

$$\omega_t \ = \ \lambda_t \ \rho_t.$$

Here, λ_t reflects a trial-level belief updating process, based on the history of stop signal delays (SSD) encountered, and is an internal noisy estimate of the prospective anticipated SSD. The parameter μ ($0 < \mu < 1$) reflects persistence in belief updating, with high persistence implying a lower decay rate of older SSDs encountered. Further, letting SSD_t be the SSD and $\mathbb{I}_{S,t}$ be a stop trial indicator,

$$\lambda_t \ = \ \frac{\left(\sum_{i=1}^{t-1} \mu^{t-i-1} \ \text{SSD}_i \ \mathbb{I}_{S,i}\right)}{\left(\sum_{i=1}^{t-1} \mu^{t-i-1} \ \ \mathbb{I}_{S,i}\right)}.$$

ρ_t is a binary variable representing cognitive state. ρ_t indicates the presence ($\rho_t = 1$) or absence ($\rho_t = 0$) of a proactive cognitive state on trial t. The proactive delayed responding is only initiated on proactive cognitive states. Proactive cognitive states are governed by a baseline proclivity for proactivity (θ_0), and a performance-monitoring based modulation (θ_1).

$$\rho_t \sim Bernoulli \left(\frac{1}{1 + \ e^{-\left(\theta_0 + \theta_1 \left(\epsilon_{G,t-1}\right) - \theta_1 \left(\epsilon_{S,t-1}\right)\right)}} \right).$$

Here, $\epsilon_{G,t-1}$ is an indicator of a go-omission (incorrectly stopping on a go trial) on the previous trial, and $\epsilon_{S,t-1}$ is an indicator of a stopping error (not stopping on a stop trial). The PRAD model assumes that the correction in terms of increasing or decreasing the probability of a proactive cognitive state following these two types of trials will be in opposite directions. The sign of θ_1 is an indicator of adaptivity or maladaptivity of the performance monitoring mechanism, and the absolute value of θ_1 denotes the sensitivity of the state-switching mechanism to errors. Probability of proactivity was the posterior mean of ρ_t, i.e., the posterior probability of $\rho_t = 1$. Proactive delaying was the posterior mean of λ_t on trials with probability of proactivity greater than 0.5, otherwise 0.

Thus, the overall effective non-decision time, as compared to traditional models, will be:

$$\tau_G \ + \ \omega_t$$

Here, τ_G is the fixed component, or the core non-decision time, while ω_t measures the strategic adjustment to the non-decision time. Note that because of this mechanism, τ_G cannot directly be compared to non-decision times from traditional models (see **SI text** for details).

(b) The stop process is modeled as a drift diffusion process with a trial-invariant non-decision time (τ_S), decision threshold (α_S), and drift rate (δ_S), but a trial-varying bias ($\beta_{S,t}$). The stop process begins at the onset of the stop signal. The initial bias is

$$\beta_{S,t} = \left(\frac{1}{1 + e^{-\gamma_1(nSSD_t - \gamma_0)}}\right).$$

Here, $nSSD_t$ reflects the number of trials since a stop signal was last encountered. This reflects an attentional mechanism that modulates the stopping bias $\beta_{S,t}$ (which varies from 0 to 1). Positive values of γ_1 result in an increase in stopping bias as $nSSD_t$ increases and vice versa. Similarly, negative values of γ_1 result in a decrease in stopping bias as $nSSD_t$ increases and vice versa. The absolute value of γ_1 measures the sensitivity to attentional modulation. The γ_0 parameter is a measure of the value of $nSSD_t$ when stopping bias is neutral (0.5).

Both the go and stop processes are implemented within a hierarchical Bayesian modeling framework in JAGS[118], using the Wiener distribution[119], which produces a joint distribution of the reaction times and the decision choice on each trial. The reaction times of the go process correspond to the reaction times for pressing the left or right buttons in response to the go stimulus. The reaction times of the stop process correspond to the SSRT. The stop process is only initiated on stop trials after the appearance of the stop stimulus (which appears after a delay corresponding to the SSD). The SSRT is not manifested as a behavioral action. Rather, if the SSRT, which is the duration of the stop process, plus the SSD on a stop trial is smaller than the go process reaction time, then the go action can be successfully inhibited (successful stop). The interaction of the basic go and stop processes can be influenced by the dynamics of the proactive delayed responding as well as the dynamics of the attentional modulation of reactive stopping. The PRAD model enables obtaining the full posterior distributions of SSRT, proactive delay in responding of the go process, and the probability of proactive cognitive states at a trial level. For further details of the model, see ref. [120]. The models were implemented in a hierarchical Bayesian framework in JAGS[118] which implements a Gibbs sampler for Markov Chain Monte Carlo (MCMC) simulations. The sampling hyperparameters, Bayesian priors, and additional computational details are described in the **SI text** and **SI Table S20**, and related control analysis is presented in **SI Table S21**.

Control Models

RVM: Random variability model is a simplification of the PRAD model without the dynamic hierarchical components and is equivalent to a full Bayesian implementation of the traditional horse-race model[117], but with the addition of allowing SSRT to vary randomly across trials, based on the parameters of a stopping drift diffusion process. It can be considered a nested version of PRAD with the following constraints applied, plus a change in some priors (**SI table S20):** Constraints: $\alpha_{G1} = \alpha_{G2}$, $\delta_{G,t} = \delta_0$ $\mathbb{S}_{LR,t}$, $\lambda_t = 0$, $\rho_t = 0$, $\beta_{S,t} = 0.5$.

FSM: Fixed SSRT model is a further simplification of RVM, where the stopping process is not explicitly modeled, but a constant SSRT value is inferred for each individual which applies to all trials.

Parameter Recovery

To test parameter recovery, we sampled 750 combinations of parameters inferred from actual data, then generated new simulated data (RT, stopping success, choice accuracy, at a trial level)

using the PRAD model and these combinations of parameters. Finally, we fit this simulated data using the PRAD model and compared the inferred parameters (recovered) to the parameters used to simulate the data. Parameter recovery is not an assessment of the validity of the PRAD model or its assumptions, nor a measure of effectiveness of the Bayesian methods used to make inferences. It does provide a way to check the implementation of the model, diagnose any potential identifiability issues [121], and understand the adequacy of the ABCD SST experimental designs for making useful model inferences.

Factor Analysis

Factor analysis was carried out using the lavaan[122] package in R[123] with the following choice of SEM (structural equation model) hyperparameters: (rotation = "oblimin", estimator = "ML", likelihood = "normal", auto.var = TRUE, auto.efa = TRUE). To compare and evaluate the adequacy of factor analysis models, we used the following criteria: CFI (comparative fit index; threshold 0.95), TLI (Tucker-Lewis index; threshold 0.95), and RMSEA (threshold < 0.08).

Analysis of bias in traditional SSRT measures versus PRAD

For each of the below, participants were split into groups based on the relevant measure, and the difference between the PRAD model inferred SSRT and non-parametric integrated SSRT (iSSRT) was assessed for group differences. Since the latter is a single value at the individual level, while the PRAD model yields a posterior distribution of SSRTs for each trial, we computed the median of the posterior distributions for each trial and used the average of these across trials to assess individual level SSRTs.

RT skewness and RT slowdown: Participants were split into four groups (quantiles) based on these measures and assessed for group differences using ANOVA analysis. Significant group differences *and* increasing SSRT difference as skewness and slowdown increased (previously reported conditions where iSSRT has been shown to underestimate SSRT[28,69]) were indicative of debiased PRAD SSRT.

Stop success rates: Participants were split into three groups based on stop success being < =25%, 25%-75% and >=75% and assessed for group differences using ANOVA analysis. Previous literature has recommended assessing non-parametric SSRT only when stopping success ranges in the 25%-75% range, with underestimation reported for high success rates [28,69]. Significant group differences *and* higher SSRT difference for high stop success rates were indicative of debiased PRAD SSRT.

Go omission rates: Participants were split into two groups based on go-omission rates being < or > 20% and assessed for group differences using a t-test. Previous literature has recommended avoiding non-parametric SSRT when go-omission rates are high[28,69]. Significant group differences *and* higher SSRT difference for high go-omission rates were indicative of debiased PRAD SSRT.

Violators (NDAR classification): Participants were split into two groups based on whether they were classified as violators as per the ABCD NDAR classification, which infers violators as children who seem to violate traditional assumptions of context independence, which are necessary for accurate non-parametric SSRT estimation. These groups were assessed for group

differences using a t-test. Significant group differences *and* higher SSRT difference for violators were indicative of debiased PRAD SSRT.

Other measures (exploratory): Other behavioral measures were used to split individuals into four groups (quantiles) and assess whether there were group differences *and* the SSRT differences across groups increased or decreased monotonically. This was done by splitting individuals based on RT SD, RT kurtosis, SSRT SD, and within individual correlation between SFR and SSD. Previous literature has not adequately assessed whether these variables may affect estimation of SSRT, and this provides an exploratory analysis, providing insights into which variables may possibly systematically bias traditional estimates.

Assessing nonergodicity

Nonergodicity in the behavioral sciences has been assessed in terms of differences in inferences that can be made about the statistics of a variable or association between multiple variables, when inferred based on within-individual analysis versus between-individual analysis [86,87]. Acknowledging that there is a strict statistical definition of nonergodicity, we rely on this intuitive understanding of nonergodicity as it has been adopted in the behavioral sciences. Thus, nonergodicity is assessed by comparing the correlation between pairs of observed or latent measures that may vary across trials: (a) Within-individual correlations assess the correlation across all trials for a single individual. We then assess the average of these correlations across individuals, and what % of individuals show +ve vs -ve correlations. (b) Between-individual correlations first summarize the trial level measures within each individual (mean or median) and then correlate the summarized measures across individuals. When the direction of the between-subject correlations is different from the average of the within-individual correlations, or more than half the individuals show within-individual correlations in the opposite direction to the between-subject correlations, we assess these relationships as characterizing nonergodic processes.

Predicting NIH Cognitive Toolbox – SVM

For predicting NIH cognitive toolbox scores, we implemented a support vector machine (SVM) regression model with a Gaussian kernel, and assessed the fit (using all the data) as well as the cross-validated prediction (using 10-fold CV), for the following uncorrected scores[124,125]: Overall NIH cognitive toolbox score, Fluid sub-score, Crystallized sub-score, and individual scores on the following tasks: dimensional change card sort test (DCCS), picture vocabulary test (PVT), list sorting working memory test (LSWM), flanker inhibitory control and attention test (Flanker), pattern comparison processing speed test (PCPST), and oral reading recognition test (ORRT). Brief details of these tasks are included in the **SI Text**. To assess the quality of predictions, we evaluated the adjusted R^2 and the correlations between predictions and actuals.

Statistical Testing

Unless stated otherwise, correlations refer to Pearson's linear correlation coefficient, p-values are based on two-tailed tests.

Software

Data were processed and analyzed using R[123], MATLAB, and JAGS[118] (version 4.3.0).

Acknowledgements.
This work was supported by the National Institutes of Health (MH121069, MH124816), the National Science Foundation (2024856), and the Stanford Maternal and Child Health Research Institute.

Competing interests
Authors declare that they have no competing interests.

Data availability
Data used in this study were from the ABCD study (https://abcdstudy.org/), held in the National Institute of Mental Health Data Archive. These data are publicly shared with eligible researchers.

Code availability
All code used in this study will be made available upon publication (https://github.com/scsnl).

References

1 Aron, A. R. From reactive to proactive and selective control: developing a richer model for stopping inappropriate responses. *Biol Psychiatry* **69**, e55-68, doi:10.1016/j.biopsych.2010.07.024 (2011).

2 Braver, T. S. The variable nature of cognitive control: a dual mechanisms framework. *Trends Cogn Sci* **16**, 106-113, doi:10.1016/j.tics.2011.12.010 (2012).

3 Braver, T. S., Paxton, J. L., Locke, H. S. & Barch, D. M. Flexible neural mechanisms of cognitive control within human prefrontal cortex. *Proceedings of the National Academy of Sciences* **106**, 7351-7356 (2009).

4 Menon, V., Adleman, N. E., White, C. D., Glover, G. H. & Reiss, A. L. Error-related brain activation during a Go/NoGo response inhibition task. *Human brain mapping* **12**, 131-143 (2001).

5 Cai, W. *et al.* Causal interactions within a frontal-cingulate-parietal network during cognitive control: Convergent evidence from a multisite–multitask investigation. *Cerebral cortex* **26**, 2140-2153 (2016).

6 Nigg, J. On the relations among self-regulation, self-control, executive functioning, effortful control, cognitive control, impulsivity, risk-taking, and inhibition for developmental psychopathology. *Journal of Child Psychology and Psychiatry* **58**, 361-383 (2017).

7 Casey, B. J., Tottenham, N., Liston, C. & Durston, S. Imaging the developing brain: what have we learned about cognitive development? *Trends in cognitive sciences* **9**, 104-110 (2005).

8 Toga, A. W., Thompson, P. M. & Sowell, E. R. Mapping brain maturation. *Trends in neurosciences* **29**, 148-159 (2006).

9 Menon, V. Developmental pathways to functional brain networks: emerging principles. *Trends in cognitive sciences* **17**, 627-640 (2013).

10 Menon, V. Large-scale brain networks and psychopathology: a unifying triple network model. *Trends in cognitive sciences* **15**, 483-506 (2011).

11 Murray, K. T. & Kochanska, G. Effortful control: Factor structure and relation to externalizing and internalizing behaviors. *Journal of abnormal child psychology* **30**, 503-514 (2002).

12 Korenblum, C. B., Chen, S. X., Manassis, K. & Schachar, R. J. Performance monitoring and response inhibition in anxiety disorders with and without comorbid ADHD. *Depression and Anxiety* **24**, 227-232 (2007).

13 Olvet, D. M. & Hajcak, G. The error-related negativity (ERN) and psychopathology: Toward an endophenotype. *Clinical psychology review* **28**, 1343-1354 (2008).

14 Hall, J. R., Bernat, E. M. & Patrick, C. J. Externalizing psychopathology and the error-related negativity. *Psychological science* **18**, 326-333 (2007).

15 Schachar, R. J. *et al.* Evidence for an error monitoring deficit in attention deficit hyperactivity disorder. *Journal of abnormal child psychology* **32**, 285-293 (2004).

16 Schachar, R., Tannock, R., Marriott, M. & Logan, G. Deficient inhibitory control in attention deficit hyperactivity disorder. *Journal of abnormal child psychology* **23**, 411-437 (1995).

17 Ivanov, I., Schulz, K. P., London, E. D. & Newcorn, J. H. Inhibitory control deficits in childhood and risk for substance use disorders: a review. *The American journal of drug and alcohol abuse* **34**, 239-258 (2008).
18 Oosterlaan, J. & Sergeant, J. A. Inhibition in ADHD, aggressive, and anxious children: A biologically based model of child psychopathology. *Journal of abnormal child psychology* **24**, 19-36 (1996).
19 Casey, B. *et al.* The adolescent brain cognitive development (ABCD) study: imaging acquisition across 21 sites. *Developmental cognitive neuroscience* **32**, 43-54 (2018).
20 Schachar, R. & Logan, G. D. Impulsivity and inhibitory control in normal development and childhood psychopathology. *Developmental psychology* **26**, 710 (1990).
21 Bonham, M. D., Shanley, D. C., Waters, A. M. & Elvin, O. M. Inhibitory control deficits in children with oppositional defiant disorder and conduct disorder compared to attention deficit/hyperactivity disorder: A systematic review and meta-analysis. *Research on Child and Adolescent Psychopathology* **49**, 39-62 (2021).
22 Alderson, R. M., Rapport, M. D. & Kofler, M. J. Attention-deficit/hyperactivity disorder and behavioral inhibition: a meta-analytic review of the stop-signal paradigm. *Journal of abnormal child psychology* **35**, 745-758 (2007).
23 McTeague, L. M. *et al.* Identification of common neural circuit disruptions in cognitive control across psychiatric disorders. *American Journal of Psychiatry* **174**, 676-685 (2017).
24 Lipszyc, J. & Schachar, R. Inhibitory control and psychopathology: a meta-analysis of studies using the stop signal task. *Journal of the International Neuropsychological Society* **16**, 1064-1076 (2010).
25 Smith, J. L., Mattick, R. P., Jamadar, S. D. & Iredale, J. M. Deficits in behavioural inhibition in substance abuse and addiction: a meta-analysis. *Drug and alcohol dependence* **145**, 1-33 (2014).
26 Lavagnino, L., Arnone, D., Cao, B., Soares, J. C. & Selvaraj, S. Inhibitory control in obesity and binge eating disorder: A systematic review and meta-analysis of neurocognitive and neuroimaging studies. *Neuroscience & Biobehavioral Reviews* **68**, 714-726 (2016).
27 Verbruggen, F. & Logan, G. D. Response inhibition in the stop-signal paradigm. *Trends in cognitive sciences* **12**, 418-424 (2008).
28 Verbruggen, F. *et al.* A consensus guide to capturing the ability to inhibit actions and impulsive behaviors in the stop-signal task. *elife* **8**, e46323 (2019).
29 Liu, S., Heitz, R. P. & Bradberry, C. W. A touch screen based stop signal response task in rhesus monkeys for studying impulsivity associated with chronic cocaine self-administration. *Journal of neuroscience methods* **177**, 67-72 (2009).
30 Godlove, D. C. *et al.* Event-related potentials elicited by errors during the stop-signal task. I. Macaque monkeys. *Journal of Neuroscience* **31**, 15640-15649 (2011).
31 Ghasemian, S., Vardanjani, M. M., Sheibani, V. & Mansouri, F. A. Dimensional bias and adaptive adjustments in inhibitory control of monkeys. *Animal Cognition* **24**, 815-828 (2021).
32 Bari, A. *et al.* Prefrontal and monoaminergic contributions to stop-signal task performance in rats. *Journal of Neuroscience* **31**, 9254-9263 (2011).

33 Caglayan, A., Stumpenhorst, K. & Winter, Y. The Stop Signal Task for Measuring Behavioral Inhibition in Mice With Increased Sensitivity and High-Throughput Operation. *Frontiers in behavioral neuroscience* **15**, 777767 (2021).
34 Eagle, D. M. *et al.* Serotonin depletion impairs waiting but not stop-signal reaction time in rats: implications for theories of the role of 5-HT in behavioral inhibition. *Neuropsychopharmacology* **34**, 1311-1321 (2009).
35 Mayse, J. D., Nelson, G. M., Park, P., Gallagher, M. & Lin, S.-C. Proactive and reactive inhibitory control in rats. *Frontiers in neuroscience* **8**, 104 (2014).
36 Carver, A., Livesey, D. & Charles, M. Age related changes in inhibitory control as measured by stop signal task performance. *International Journal of Neuroscience* **107**, 43-61 (2001).
37 Cai, W. *et al.* Hyperdirect insula-basal-ganglia pathway and adult-like maturity of global brain responses predict inhibitory control in children. *Nature communications* **10**, 4798 (2019).
38 Williams, B. R., Ponesse, J. S., Schachar, R. J., Logan, G. D. & Tannock, R. Development of inhibitory control across the life span. *Developmental psychology* **35**, 205 (1999).
39 Tsvetanov, K. A. *et al.* Activity and connectivity differences underlying inhibitory control across the adult life span. *Journal of Neuroscience* **38**, 7887-7900 (2018).
40 Hsu, H. M. & Hsieh, S. Age-related post-error slowing and stimulus repetition effect in motor inhibition during a stop-signal task. *Psychological research* **86**, 1108-1121 (2022).
41 Soreni, N., Crosbie, J., Ickowicz, A. & Schachar, R. Stop signal and conners' continuous performance tasks: Test—retest reliability of two inhibition measures in adhd children. *Journal of Attention Disorders* **13**, 137-143 (2009).
42 Senderecka, M., Grabowska, A., Szewczyk, J., Gerc, K. & Chmylak, R. Response inhibition of children with ADHD in the stop-signal task: An event-related potential study. *International Journal of Psychophysiology* **85**, 93-105 (2012).
43 Adams, N. C. & Jarrold, C. Inhibition in autism: Children with autism have difficulty inhibiting irrelevant distractors but not prepotent responses. *Journal of autism and developmental disorders* **42**, 1052-1063 (2012).
44 Schmitt, L. M., White, S. P., Cook, E. H., Sweeney, J. A. & Mosconi, M. W. Cognitive mechanisms of inhibitory control deficits in autism spectrum disorder. *Journal of Child Psychology and Psychiatry* **59**, 586-595 (2018).
45 Geurts, H. M., van den Bergh, S. F. & Ruzzano, L. Prepotent response inhibition and interference control in autism spectrum disorders: Two meta-analyses. *Autism Research* **7**, 407-420 (2014).
46 Li, F.-F., Chen, X.-l., Zhang, Y.-t., Li, R.-t. & Li, X. The role of prepotent response inhibition and interference control in depression. *Cognitive Neuropsychiatry* **26**, 441-454 (2021).
47 Yu, F. *et al.* Repetitive transcranial magnetic stimulation promotes response inhibition in patients with major depression during the stop-signal task. *Journal of Psychiatric Research* (2022).
48 Sjoerds, Z., Van Den Brink, W., Beekman, A., Penninx, B. & Veltman, D. Response inhibition in alcohol-dependent patients and patients with depression/anxiety: a functional magnetic resonance imaging study. *Psychological medicine* **44**, 1713-1725 (2014).

49 Kalanthroff, E. *et al.* The role of response inhibition in medicated and unmedicated obsessive-compulsive disorder patients: evidence from the stop-signal task. *Depression and anxiety* **34**, 301-306 (2017).
50 Sarkar, S. *et al.* Effects of diazepam on reaction times to stop and go. *Frontiers in human neuroscience* **14**, 567177 (2020).
51 Swann, N. *et al.* Deep brain stimulation of the subthalamic nucleus alters the cortical profile of response inhibition in the beta frequency band: a scalp EEG study in Parkinson's disease. *Journal of Neuroscience* **31**, 5721-5729 (2011).
52 Cai, W., George, J. S., Verbruggen, F., Chambers, C. D. & Aron, A. R. The role of the right presupplementary motor area in stopping action: two studies with event-related transcranial magnetic stimulation. *Journal of neurophysiology* **108**, 380-389 (2012).
53 Wessel, J. R., Diesburg, D. A., Chalkley, N. H. & Greenlee, J. D. A causal role for the human subthalamic nucleus in non-selective cortico-motor inhibition. *Current biology* **32**, 3785-3791. e3783 (2022).
54 Li, C.-s. R., Milivojevic, V., Kemp, K., Hong, K. & Sinha, R. Performance monitoring and stop signal inhibition in abstinent patients with cocaine dependence. *Drug and alcohol dependence* **85**, 205-212 (2006).
55 Li, C.-s. R. *et al.* Neural correlates of impulse control during stop signal inhibition in cocaine-dependent men. *Neuropsychopharmacology* **33**, 1798-1806 (2008).
56 Zhao, W. *et al.* Response inhibition deficits in insomnia disorder: an event-related potential study with the stop-signal task. *Frontiers in Neurology* **9**, 610 (2018).
57 Covassin, N. *et al.* Cognitive performance and cardiovascular markers of hyperarousal in primary insomnia. *International journal of psychophysiology* **80**, 79-86 (2011).
58 Schmid, J. M., Labuhn, A. S. & Hasselhorn, M. Response inhibition and its relationship to phonological processing in children with and without dyslexia. *International Journal of Disability, Development and Education* **58**, 19-32 (2011).
59 De Weerdt, F., Desoete, A. & Roeyers, H. Behavioral inhibition in children with learning disabilities. *Research in developmental disabilities* **34**, 1998-2007 (2013).
60 Svaldi, J., Naumann, E., Trentowska, M. & Schmitz, F. General and food-specific inhibitory deficits in binge eating disorder. *International Journal of Eating Disorders* **47**, 534-542 (2014).
61 Fiterman, O. & Raz, S. Cognitive, neural and endocrine functioning during late pregnancy: An Event-Related Potentials study. *Hormones and behavior* **116**, 104575 (2019).
62 Barnes, J. J., Dean, A. J., Nandam, L. S., O'Connell, R. G. & Bellgrove, M. A. The molecular genetics of executive function: role of monoamine system genes. *Biological psychiatry* **69**, e127-e143 (2011).
63 Verbruggen, F. & Logan, G. D. Models of response inhibition in the stop-signal and stop-change paradigms. *Neuroscience & Biobehavioral Reviews* **33**, 647-661 (2009).
64 Logan, G. D. & Cowan, W. B. On the ability to inhibit thought and action: A theory of an act of control. *Psychological review* **91**, 295 (1984).
65 Logan, G. D., 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. *Psychological review* **121**, 66 (2014).
66 Bissett, P. G., Jones, H. M., Poldrack, R. A. & Logan, G. D. Severe violations of independence in response inhibition tasks. *Science advances* **7**, eabf4355 (2021).

67 Bissett, P. G., Hagen, M. P., Jones, H. M. & Poldrack, R. A. Design issues and solutions for stop-signal data from the Adolescent Brain Cognitive Development (ABCD) study. *Elife* **10**, e60185 (2021).
68 Bissett, P. G. *Evaluating the independent race model for the stop signal paradigm: Context independence is violated at short stop signal delays*. (Vanderbilt University, 2014).
69 Verbruggen, F., Chambers, C. D. & Logan, G. D. Fictitious inhibitory differences: how skewness and slowing distort the estimation of stopping latencies. *Psychological science* **24**, 352-362 (2013).
70 Weigard, A., Matzke, D., Tanis, C. & Heathcote, A. A cognitive process modeling framework for the ABCD study stop-signal task. *Developmental Cognitive Neuroscience* **59**, 101191 (2023).
71 Weigard, A., Matzke, D., Tanis, C. & Heathcote, A. Cognitive process modeling addresses context independence violations in the ABCD Study stop-signal task. *bioRxiv*, 2021.2007. 2026.453872 (2021).
72 Doekemeijer, R. A., Dewulf, A., Verbruggen, F. & Boehler, C. N. Proactively adjusting stopping: response inhibition is faster when stopping occurs frequently. *Journal of cognition* **6** (2023).
73 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. *Clinical Psychological Science* **7**, 856-872 (2019).
74 Cai, W., Oldenkamp, C. L. & Aron, A. R. A proactive mechanism for selective suppression of response tendencies. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **31**, 5965-5969, doi:10.1523/JNEUROSCI.6292-10.2011 (2011).
75 Aron, A. R. From reactive to proactive and selective control: developing a richer model for stopping inappropriate responses. *Biological psychiatry* **69**, e55-e68 (2011).
76 Elchlepp, H., Lavric, A., Chambers, C. D. & Verbruggen, F. Proactive inhibitory control: A general biasing account. *Cognitive Psychology* **86**, 27-61 (2016).
77 Verbruggen, F. & Logan, G. D. Proactive adjustments of response strategies in the stop-signal paradigm. *Journal of Experimental Psychology: Human Perception and Performance* **35**, 835 (2009).
78 Vink, M., Kaldewaij, R., Zandbelt, B. B., Pas, P. & du Plessis, S. The role of stop-signal probability and expectation in proactive inhibition. *European Journal of Neuroscience* **41**, 1086-1094 (2015).
79 Chikazoe, J. *et al.* Preparation to inhibit a response complements response inhibition during performance of a stop-signal task. *Journal of Neuroscience* **29**, 15870-15877 (2009).
80 Swann, N. C. *et al.* Roles for the pre-supplementary motor area and the right inferior frontal gyrus in stopping action: electrophysiological responses and functional and structural connectivity. *Neuroimage* **59**, 2860-2870 (2012).
81 Cai, W. *et al.* Both reactive and proactive control are deficient in children with ADHD and predictive of clinical symptoms. *Translational psychiatry* **13**, 179 (2023).

82 Ide, J. S., Shenoy, P., Angela, J. Y. & Chiang-Shan, R. L. Bayesian prediction and evaluation in the anterior cingulate cortex. *Journal of Neuroscience* **33**, 2039-2047 (2013).
83 Cai, W., Chen, T., Ide, J. S., Li, C.-S. R. & Menon, V. Dissociable fronto-operculum-insula control signals for anticipation and detection of inhibitory sensory cue. *Cerebral cortex* **27**, 4073-4082 (2017).
84 Shenoy, P. & Yu, A. J. Rational decision-making in inhibitory control. *Frontiers Human Neuroscience* **5**, 48 (2011).
85 Medaglia, J. D., Ramanathan, D. M., Venkatesan, U. M. & Hillary, F. G. The challenge of non-ergodicity in network neuroscience. *Network: Computation in Neural Systems* **22**, 148-153 (2011).
86 Mangalam, M. & Kelty-Stephen, D. G. Point estimates, Simpson's paradox, and nonergodicity in biological sciences. *Neuroscience & Biobehavioral Reviews* **125**, 98-107 (2021).
87 Molenaar, P. C. On the implications of the classical ergodic theorems: Analysis of developmental processes has to focus on intra-individual variation. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology* **50**, 60-69 (2008).
88 Meyer-Lindenberg, A. The non-ergodic nature of mental health and psychiatric disorders: implications for biomarker and diagnostic research. *World Psychiatry* **22**, 272 (2023).
89 Ode, S., Robinson, M. D. & Hanson, D. M. Cognitive-emotional dysfunction among noisy minds: predictions from individual differences in reaction time variability. *Cognition and Emotion* **25**, 307-327 (2011).
90 Kofler, M. J. *et al.* Reaction time variability in ADHD: a meta-analytic review of 319 studies. *Clinical psychology review* **33**, 795-811 (2013).
91 Karalunas, S. L., Geurts, H. M., Konrad, K., Bender, S. & Nigg, J. T. Annual research review: Reaction time variability in ADHD and autism spectrum disorders: Measurement and mechanisms of a proposed trans-diagnostic phenotype. *Journal of Child Psychology and Psychiatry* **55**, 685-710 (2014).
92 Kuntsi, J. Commentary: From noise to insight? Reaction time variability in ADHD and autism spectrum disorders–a commentary on Karalunas et al.(2014). *Journal of Child Psychology and Psychiatry* **55**, 711-713 (2014).
93 Ma, N. & Yu, A. J. Statistical learning and adaptive decision-making underlie human response time variability in inhibitory control. *Frontiers in psychology* **6**, 1046 (2015).
94 Tamm, L., Epstein, J. N. & Becker, S. P. A preliminary investigation of reaction time variability in relation to social functioning in children evaluated for ADHD. *Child neuropsychology* **25**, 885-898 (2019).
95 Paraskevopoulou, S. E., Coon, W. G., Brunner, P., Miller, K. J. & Schalk, G. Within-subject reaction time variability: Role of cortical networks and underlying neurophysiological mechanisms. *Neuroimage* **237**, 118127 (2021).
96 Cañigueral, R. *et al.* Adaptiveness of fluctuations in intra-individual variability of performance is process-dependent in middle childhood. (2022).
97 Epstein, J. N. *et al.* Examining reaction time variability on the stop-signal task in the ABCD study. *Journal of the International Neuropsychological Society*, 1-11 (2022).
98 Aristodemou, M., Rommelse, N. & Kievit, R. Attentiveness modulates reaction-time variability: findings from a population-based sample of 1032 children. (2022).

99 Yamashita, A. *et al.* Variable rather than extreme slow reaction times distinguish brain states during sustained attention. *Scientific reports* **11**, 1-13 (2021).
100 Rieger, M. & Gauggel, S. Inhibitory after-effects in the stop signal paradigm. *British Journal of Psychology* **90**, 509-518 (1999).
101 Matzke, D., Love, J. & Heathcote, A. A Bayesian approach for estimating the probability of trigger failures in the stop-signal paradigm. *Behavior research methods* **49**, 267-281 (2017).
102 Doekemeijer, R., Verbruggen, F. & Boehler, C. Face the (trigger) failure: Trigger failures strongly drive the effect of reward on response inhibition. *Cortex* **139**, 166-177 (2021).
103 Skippen, P. *et al.* Reliability of triggering inhibitory process is a better predictor of impulsivity than SSRT. *Acta psychologica* **192**, 104-117 (2019).
104 Perri, R. L. Is there a proactive and a reactive mechanism of inhibition? Towards an executive account of the attentional inhibitory control model. *Behavioural brain research* **377**, 112243 (2020).
105 Lee, W.-T. & Kang, M.-S. Electrophysiological evidence for distinct proactive control mechanisms in a stop-signal task: an individual differences approach. *Frontiers in Psychology* **11**, 1105 (2020).
106 Hu, S. & Li, C. S. R. Neural processes of preparatory control for stop signal inhibition. *Human brain mapping* **33**, 2785-2796 (2012).
107 Mirabella, G. Vol. 13 804 (MDPI, 2023).
108 Fisher, A. J., Medaglia, J. D. & Jeronimus, B. F. Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences* **115**, E6106-E6115 (2018).
109 Curran, P. J. & Bauer, D. J. The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual review of psychology* **62**, 583-619 (2011).
110 Molenaar, P. C. A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement* **2**, 201-218 (2004).
111 Jahfari, S., Stinear, C. M., Claffey, M., Verbruggen, F. & Aron, A. R. Responding with restraint: what are the neurocognitive mechanisms? *Journal of cognitive neuroscience* **22**, 1479-1492 (2010).
112 Castro-Meneses, L. J., Johnson, B. W. & Sowman, P. F. The effects of impulsivity and proactive inhibition on reactive inhibition and the go process: insights from vocal and manual stop signal tasks. *Frontiers in Human Neuroscience* **9**, 529 (2015).
113 Casey, B. J. *et al.* The adolescent brain cognitive development (ABCD) study: imaging acquisition across 21 sites. *Developmental cognitive neuroscience* **32**, 43-54 (2018).
114 Garavan, H. *et al.* Recruiting the ABCD sample: Design considerations and procedures. *Developmental cognitive neuroscience* **32**, 16-22 (2018).
115 Volkow, N. D. *et al.* The conception of the ABCD study: From substance use to a broad NIH collaboration. *Dev Cogn Neurosci* **32**, 4-7, doi:10.1016/j.dcn.2017.10.002 (2018).
116 Logan, G. D., Schachar, R. J. & Tannock, R. Impulsivity and inhibitory control. *Psychological science* **8**, 60-64 (1997).
117 Band, G. P., Van Der Molen, M. W. & Logan, G. D. Horse-race model simulations of the stop-signal procedure. *Acta psychologica* **112**, 105-142 (2003).
118 Plummer, M. in *Proceedings of the 3rd international workshop on distributed statistical computing.* 1-10 (Vienna, Austria).

119 Wabersich, D. & Vandekerckhove, J. Extending JAGS: A tutorial on adding custom distributions to JAGS (with a diffusion model example). *Behavior research methods* **46**, 15-28 (2014).
120 Mistry, P. K., Warren, S. L., Cai, W., Branigan, N. K. & Menon, V. Proactive, reactive, and attentional dynamics in inhibitory control. (In Prep.).
121 Matzke, D., Logan, G. D. & Heathcote, A. A cautionary note on evidence-accumulation models of response inhibition in the stop-signal paradigm. *Computational Brain & Behavior* **3**, 269-288 (2020).
122 Rosseel, Y. lavaan: An R package for structural equation modeling. *Journal of statistical software* **48**, 1-36 (2012).
123 Team, R. C. RA language and environment for statistical computing, R Foundation for Statistical. *Computing* (2020).
124 Slotkin, J. *et al.* NIH toolbox. *Technical Manual.[Google Scholar]* (2012).
125 Akshoomoff, N. *et al.* VIII. NIH Toolbox Cognition Battery (CB): composite scores of crystallized, fluid, and overall cognition. *Monographs of the Society for Research in Child Development* **78**, 119-132 (2013).

Figures

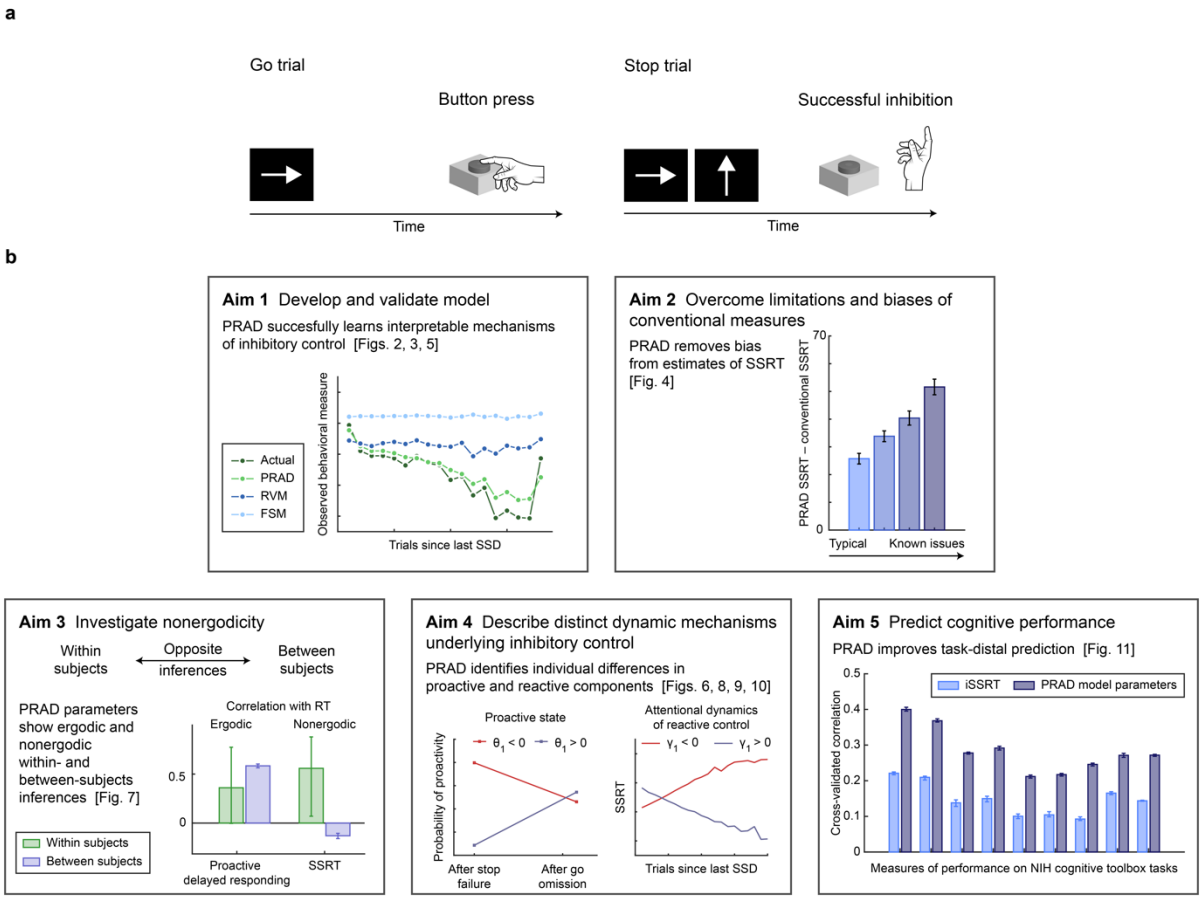

Figure 1. Overview of the stop signal task and analysis pipeline. (A) Illustration of Go and Stop trials in the stop signal task. On Go trials, participants respond to indicate arrow direction; on Stop trials, they must inhibit their response when the stop signal appears. **(B)** Schematic representation of the data analysis and modeling approach used in this study, outlining the process from raw data to model-derived insights.

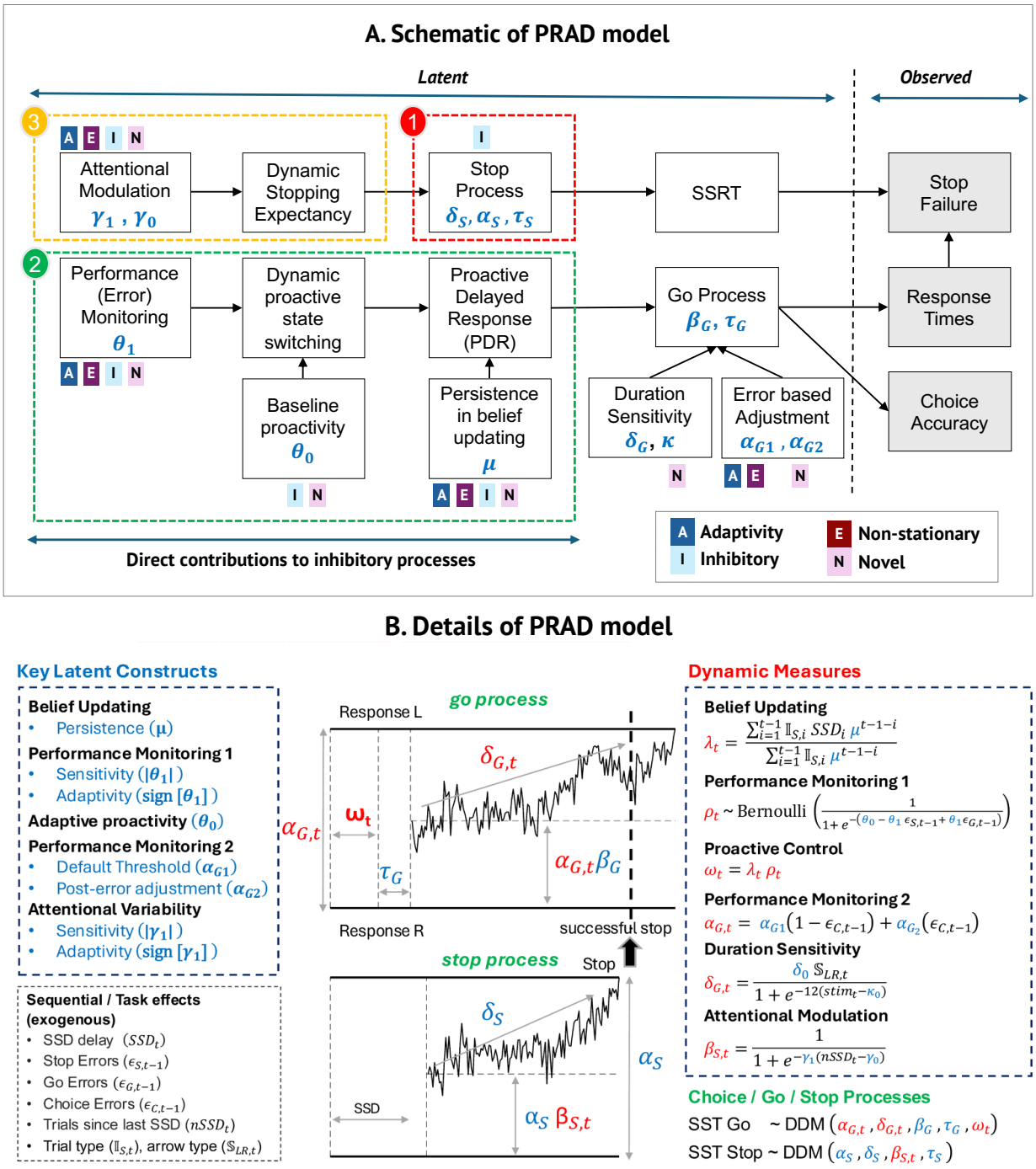

Figure 2. Schematic of the Proactive Reactive and Attentional Dynamics (PRAD) model. (A) Overview of how latent cognitive processes interact to produce observed behavior. The model infers latent variables related to three mechanisms of dynamic inhibitory control: basic reactive inhibition, proactive response delaying, and attentional modulation of stopping expectancy. **(B)** Mathematical details of the model, showing two drift diffusion processes (go, stop) augmented with dynamic parameters. Blue - individual-level parameters, Red - trial-level dynamic parameters.

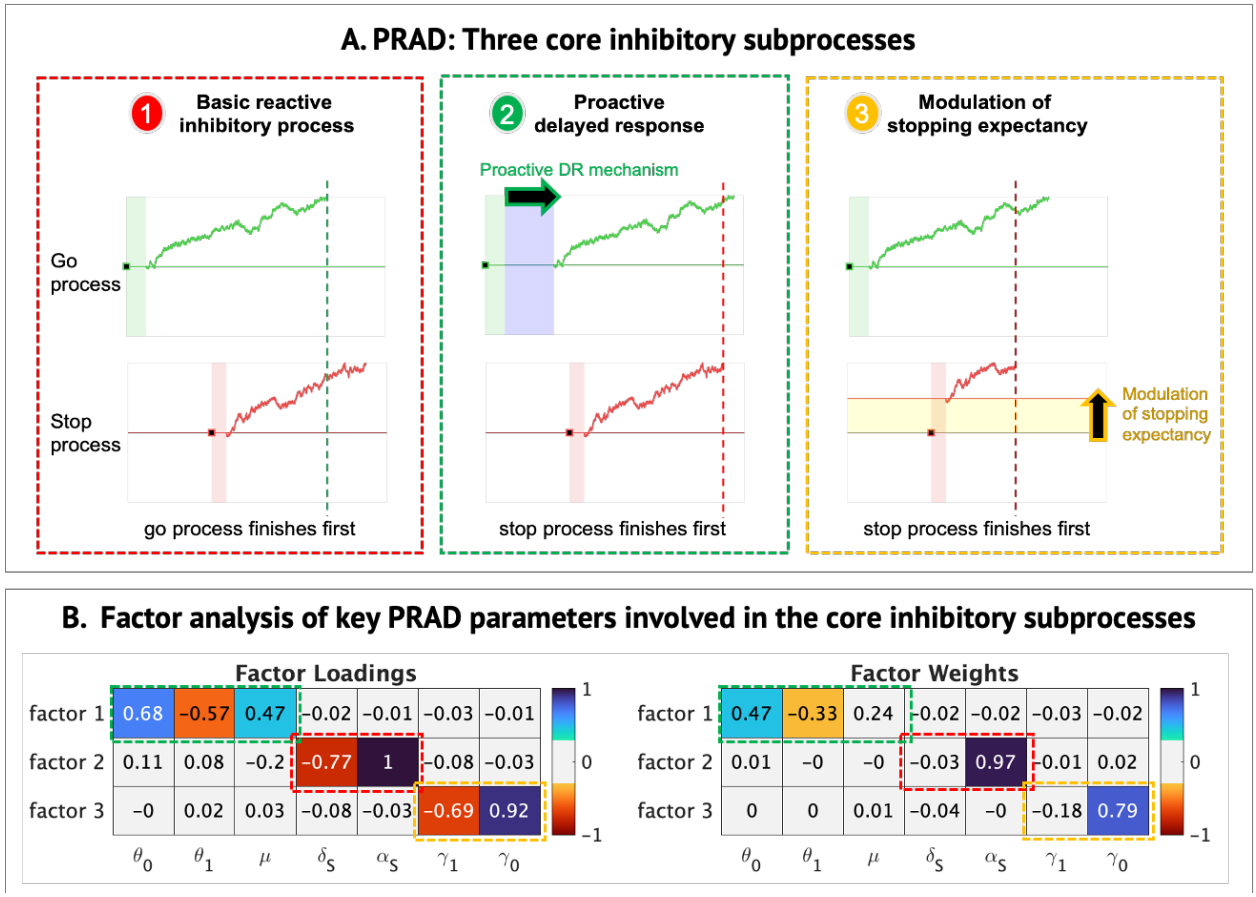

Figure 3. Core components of the PRAD model. (A) Illustration of three key inhibitory subprocesses: (1) basic reactive inhibition (red box), where competing Stop and Go processes determine trial outcome; (2) proactive delayed responding (PDR, green box), where anticipatory delaying of the Go process can lead to successful stopping; and (3) attentional modulation of stopping expectancy (yellow box), where top-down modulation can influence SSRT. **(B)** Factor analysis results showing how seven key PRAD parameters load onto three distinct factors corresponding to these subprocesses – proactive control, basic reactive control, and attentional modulation – demonstrating the model's ability to dissociate different aspects of inhibitory control.

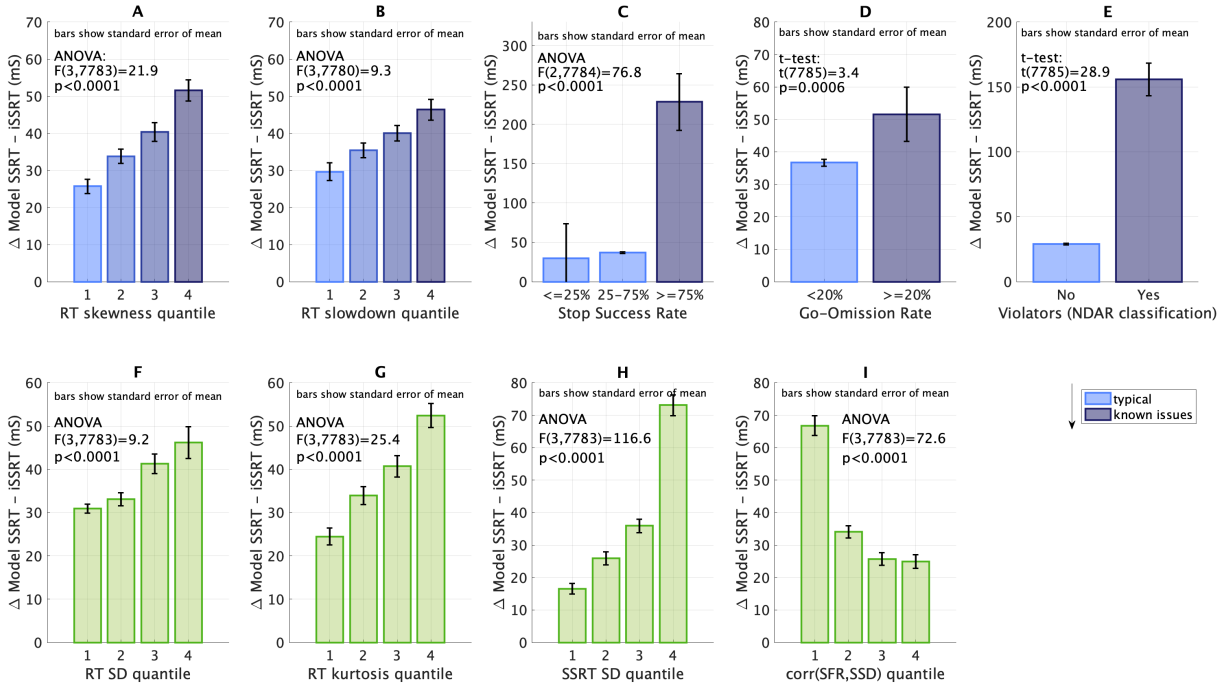

Figure 4. Debiasing of stop-signal reaction time (SSRT) estimates. Comparison of PRAD-inferred SSRT with traditional non-parametric integrated SSRT (iSSRT) across various conditions. **(A-E)** Conditions known to bias SSRT estimates, including high RT skewness, RT slowdown, stop success rates, go omission rates, and context independence violations. **(F-I)** Additional conditions revealed by PRAD, such as high RT variability and kurtosis, high SSRT variability, and low SSD-SFR correlations. Higher PRAD SSRT estimates in these conditions demonstrate the model's ability to compensate for known biases and reveal new insights into SSRT estimation.

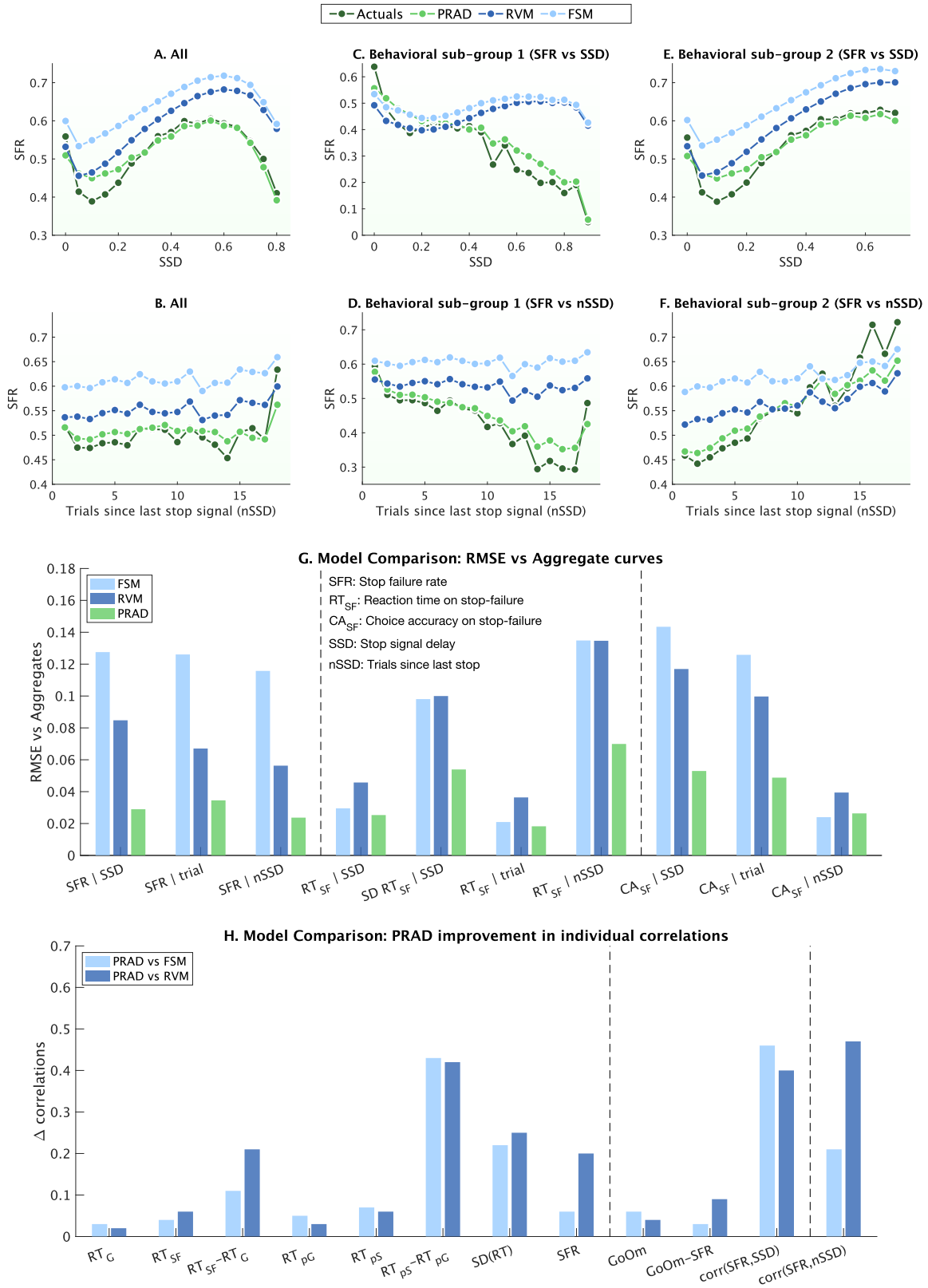

Figure 5. Model fit comparisons. (A-F) Aggregate behavior patterns for stop-failure rates (SFR) across different stop-signal delays (SSD) and number of trials since last stop signal

(nSSD), comparing actual data with predictions from PRAD and control models (RVM and FSM). PRAD uniquely accounts for subgroup differences in SFR-SSD and SFR-nSSD relationships. **(G)** Root mean square error (RMSE) comparisons across models for various behavioral measures, showing PRAD's superior fit. RMSE computed based on the distance between binned aggregate curves (actuals vs model posterior generated values). **(H)** Improvements in individual-level fit correlations using PRAD compared to control models across multiple behavioral measures. Dark green – actuals, Light green – PRAD, dark blue – RVM, light blue – FSM. RVM = Random Variability control model; FSM = Fixed SSRT control model.

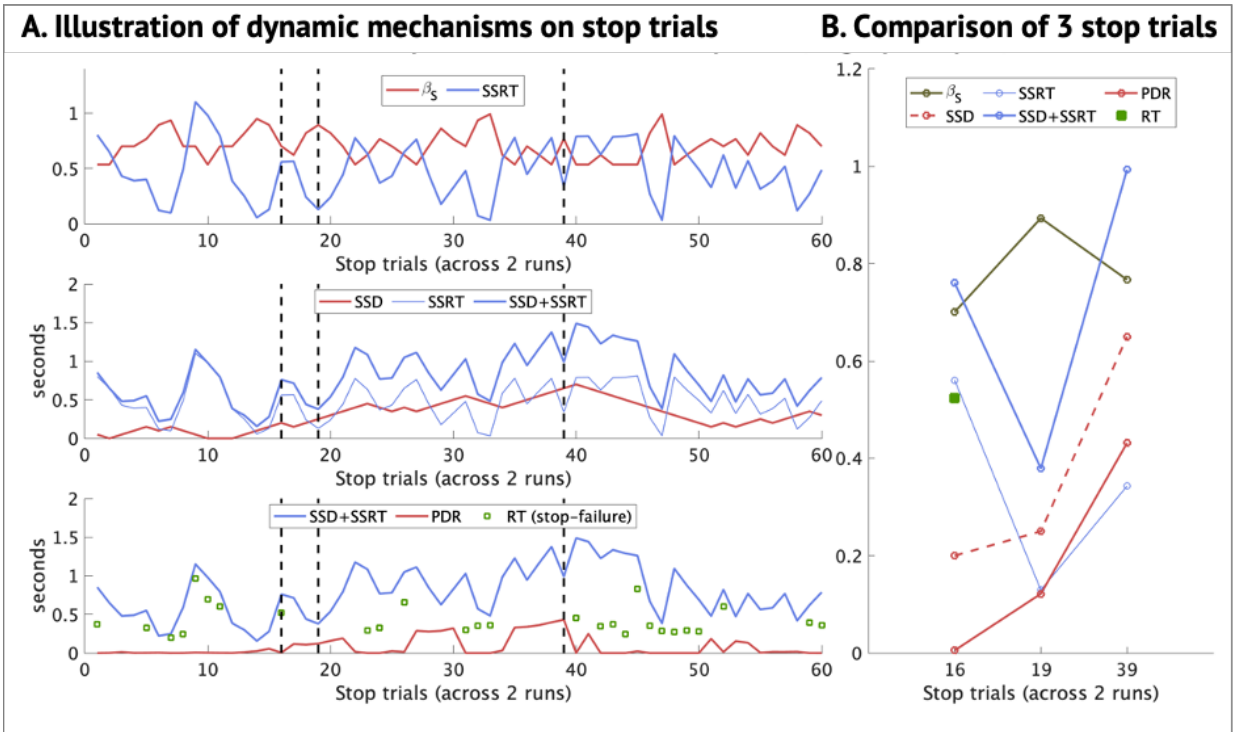

Figure 6. Illustrative single-participant dynamics on stop trials. (A) Trial-by-trial fluctuations in stopping expectancy, SSRT, SSD, PDR, and RT, illustrating the complex interplay of these variables. **(B)** Comparison of three specific stop trials (16, 19, and 39) demonstrating how different combinations of SSD, SSRT, and PDR lead to successful stops or failures, highlighting the model's ability to capture nuanced trial-level dynamics. On stop trial 16, SSD is low (easy trial), but it still results in a stop-failure as there is no proactive delaying of the go process (PDR ~ 0), and RT < (SSD+SSRT). On stop trial 19, SSD is higher (difficult trial), but SSRT is much lower because stopping expectancy is higher (attentional modulation of SSRT), resulting in a successful stop. On stop trial 39, SSD is even higher (very difficult), and stopping expectancy is not very high, resulting in a high SSRT. However, in spite of SSD+SSRT being higher than stop trial 16, this trial does not result in a stop-failure because of the influence of proactive delay of the go process (high PDR).

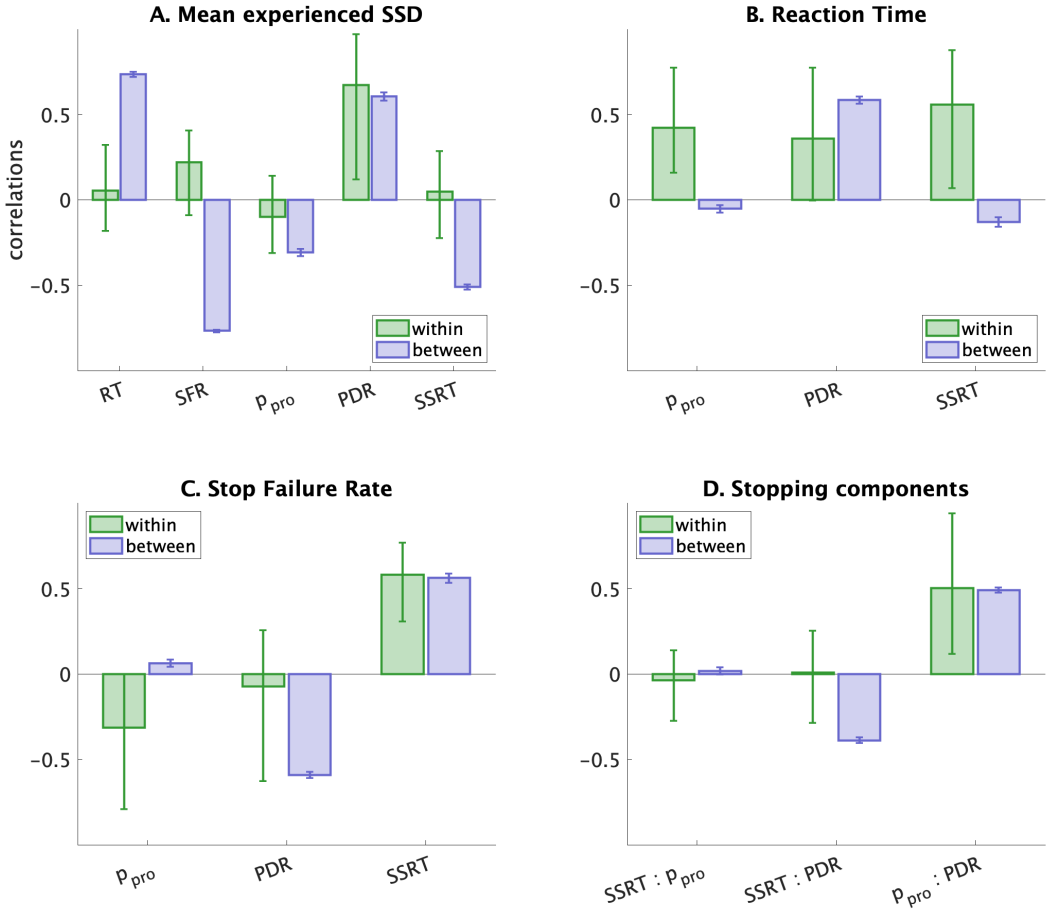

Figure 7. Evidence for nonergodicity in cognitive dynamics. Comparison of within-subject (green) and between-subject (blue) correlations for various pairs of observed and latent variables. Differences between these correlations, especially sign changes, indicate nonergodic processes. Notable nonergodic effects are observed in some of the relationships between mean experienced SSD, SFR, SSRT, RT, PDR, and probability of proactive state. The between-subjects values plotted are based on average values for individuals and bootstrapping the sample. SFR: Stop failure rate, SSRT: Stop-signal reaction time, RT: Reaction time, PDR: Proactive delayed responding, p_{pro}: probability of proactive cognitive state. Error bars reflect 2.5^{th} to 97.5^{th} percentiles for within subject correlations and 2.5^{th} to 97.5^{th} percentiles based on bootstrapped samples for between subject correltions.

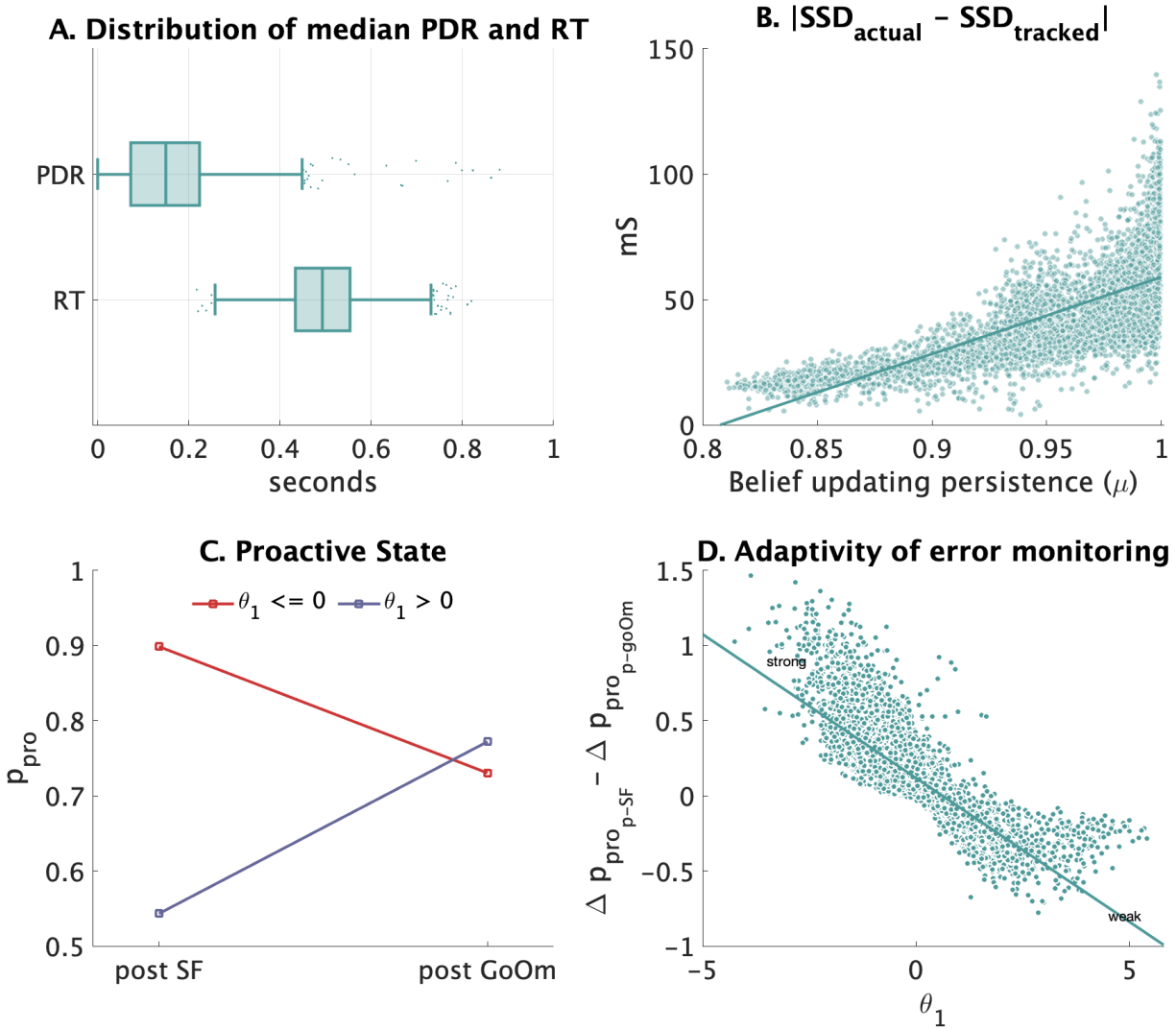

Figure 8. Proactive Delayed Response (PDR) dynamics. (A) Distribution of median PDR and RT across individuals, showing the range of proactive control strategies. **(B)** Relationship between persistence in belief updating (μ) and SSD tracking errors, illustrating maladaptive effects of high persistence. **(C)** Differences in post stop-failure proactivity between θ_1 subgroups. **(D)** Relationship between error monitoring adaptivity θ_1 and differences in proactive delaying probability following stop-failures vs. go trials.

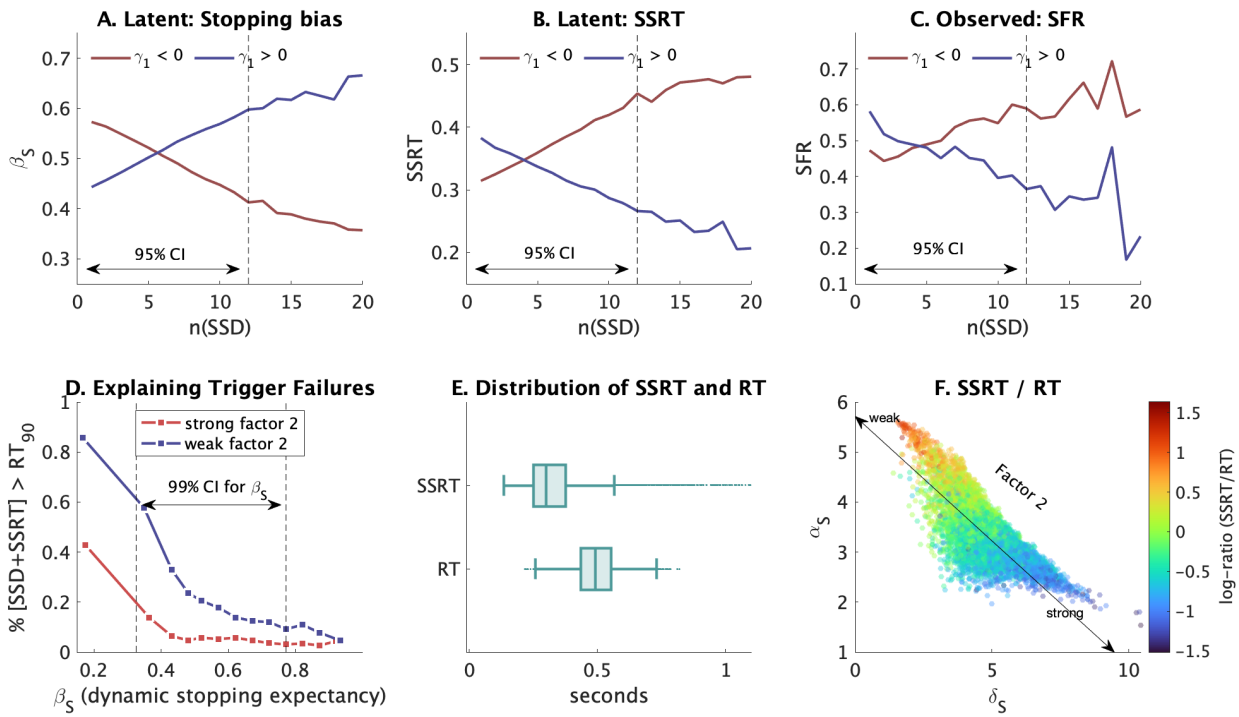

Figure 9. Attentional Modulation of Stopping (AMS) dynamics. (A-C) Evolution of stopping dynamics (expectancy, SSRT, and SFR) with increasing trials since last stop signal (nSSD) for different attentional modulation subgroups ($\gamma_1 > 0$ vs $\gamma_1 < 0$), showing adaptive and maladaptive patterns. **(D)** Relationship between stopping expectancy and proportion of high SSRT trials, differentiated by basic reactive process strength. **(E)** Distribution of median SSRT and RT across individuals. **(F)** Log-Ratio of SSRT/RT as a function of stopping process drift rate and threshold, illustrating parameter interactions.

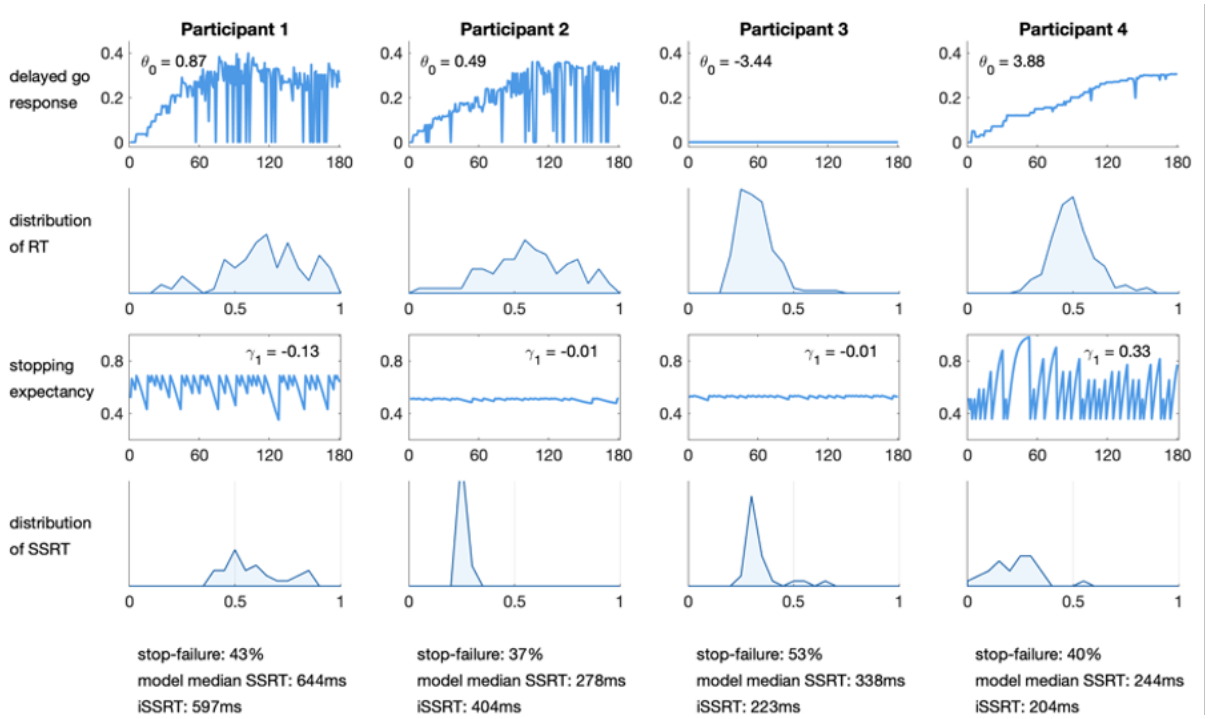

Figure 10. Illustration of variability in PDR and AMS processes across four representative participants. Comparison of trial-by-trial fluctuations in PDR, RT, stopping expectancy, and SSRT, demonstrating how similar overall performance can arise from different underlying cognitive dynamics. This figure highlights how the PRAD model can dissociate different sources of intra-individual variability that may be indistinguishable using traditional measures.

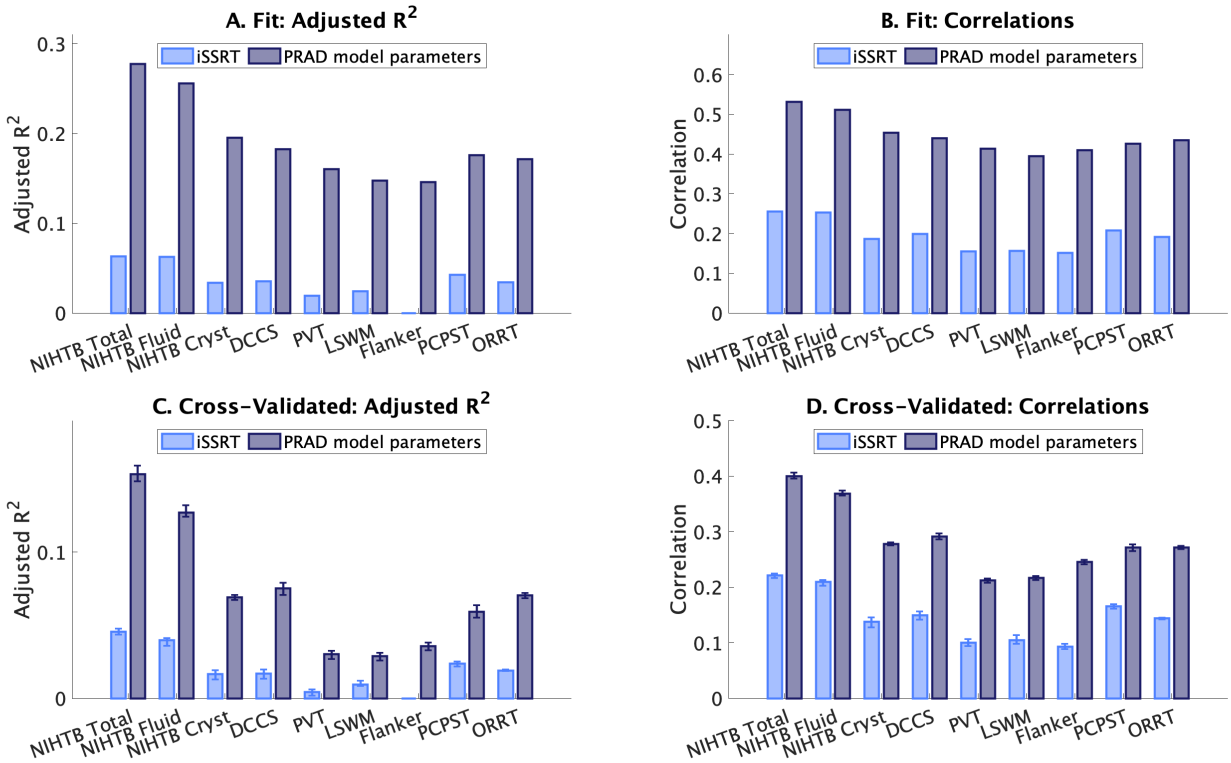

Figure 11. Predictive power of PRAD for NIH Toolbox cognitive tasks. Comparison of model fit and cross-validated performance using traditional SSRT (iSSRT) versus PRAD model parameters. Results show superior predictive power of PRAD for task-distal cognitive measures, with higher adjusted R^2 and correlations between predicted and actual values in both model fit and cross-validated analyses. Error bars show the 2.5th to 97.5th percentiles across 10-fold cross-validation.