Title
Meta-Learning of Dynamic Policy Adjustments in Inhibitory Control Tasks

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors
Chatterjee, Soumya
Kumar, Aakriti
Shenoy, Pradeep

Publication Date
2022

Peer reviewed

Meta-Learning of Dynamic Policy Adjustments in Inhibitory Control Tasks

Soumya Chatterjee* **(soumyach@google.com)**
Google Research India

Aakriti Kumar* **(aakritk@uci.edu)**
University of California, Irvine

Pradeep Shenoy (shenoypradeep@google.com)
Google Research India

Abstract

Simple perceptual decision-making tasks such as the Stroop and flanker tasks are popular as a method of measuring individual variation in the processing of conflicting visual stimuli–for instance, the difference in accuracy on stimuli with and without conflict. A major challenge in applying these tasks, for instance, to compare two different populations of subjects, is the low reliability of the nonparametric measures of performance in the tasks. Here, we model dynamic adjustments in decision policies often seen in human behavior, thereby capturing trial-by-trial variation in decision policies, in addition to the classically used average statistics. We propose a recurrent network model to capture behavioral strategies in the task in a model-agnostic manner, and to overcome small-sample learning challenges by pooling across subjects. We show that by splitting the learning into a complex, shared meta-model and simple subject-specific parameters, we learn significantly better predictive models, and also identify latent dimensions indexing the decision policy that may serve as a better measure of individual differences in the task.

Keywords: Artificial Intelligence; Cognitive Neuroscience; Decision Making; Machine Learning; Neural Networks

Introduction

Decades of research in cognitive neuroscience have focused on simple behavioral tasks thought to measure aspects of decision-making and related cognitive processes. For instance, the Eriksen flanker task (Figure 1, Eriksen and Eriksen (1974)) purportedly measures an individual's ability to suppress or 'inhibit' irrelevant visual information[1]. Such tasks are widely used as a battery of behavioral experiments in large-scale studies–for instance, in longitudinal studies of adolescent development (Garavan et al., 2018), or psychiatric conditions & clinical outcomes (Victor et al., 2018). Yet, recent research raises significant concerns about these tasks: that they do not correlate well with life outcomes (Eisenberg et al., 2019), and that indices of performance in these tasks show unacceptably high within-subject variation in repeat measurements (Enkavi et al., 2019; Hedge, Powell, & Sumner, 2018). This is true both for empirical measures such as error rates and average response times (RT) on different types of trials, and also for latent variables/parameters estimated

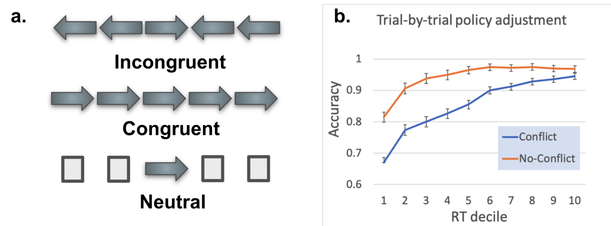


Figure 1: Sequential updating in inhibitory control tasks. (a) Eriksen Flanker task: participants respond to the direction of the central arrow while ignoring flanking stimuli. Response times (RT) and accuracy deltas between congruent & incongruent trials are considered indices of cognitive control. (b) Relation between RT decile and accuracies, pooled across subjects (deciling within subject), showing that RT variance is influenced by trial-by-trial adjustments to decision policy.

from model fits to RT/accuracy distributions using classical approaches like the race model (Ratcliff & Rouder, 1998; Logan & Cowan, 1984b).

We argue that classical empirical measures and computational models focus on average performance, and fail to capture the full range of behavior. In particular, well-known *stimulus-history-based* changes in response times and accuracy (Laming, 1968) contain valuable information about trial-by-trial adjustments to subjects' decision policies, and explain significant behavioral variability. Driven by this insight, we propose and address the novel challenge of extracting subject-specific parameters of *dynamic policy adjustments* as a more informative individual index of task performance.[2]

Our contributions are as follows: (a) We propose a deep recurrent model (RECNET) to predict sequential adjustments to speed-accuracy tradeoffs, (b) We *pool data across subjects* in a meta-learning framework to effectively learn the recurrent model of behavior, and (c) We learn a *compact parametrization* of each subject in the meta-model to customize it to each subjects' individual behavior. Our proposed meta-learning model is significantly more accurate at predicting behavior

*Equal contribution

[1] We describe only one task for illustration; there are a wide range of very well-studied tasks probing various aspects of the notion of "inhibitory control", e.g., the go-nogo task (Donders, 1969; Logan & Cowan, 1984a), the Stroop task (Stroop, 1935), the stop-signal task (Lappin & Eriksen, 1966), etc.

[2] Although much studied in recent research, from Bayes nets (Yu & Cohen, 2009) to hypernetworks (Dezfouli, Ashtiani, et al., 2019), sequential effects models have not directly addressed issues of test-retest reliability; indeed, recent proposals (Dezfouli, Ashtiani, et al., 2019; Yu & Cohen, 2009) fare poorly in our experiments on measures of reliability. See Related Work & Results sections for details.

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society.*

compared to simple pooled models, and the predictive power transfers easily to new sessions of data. In particular, we demonstrate that given a prelearned meta-model, our subject parameters can be recovered with very little data, and show significantly better test-retest reliability (Intraclass correlations (Shrout & Fleiss, 1979)) compared to classical behavioral and model-based measures.

Our focus in this work is primarily on *reliable estimation of subject-specific performance indices* as a machine learning challenge; making connections to cognitive processes or proposing new theories is beyond the scope of this paper. Nevertheless, we believe that our work (a) addresses a key challenge for the broader field of cognitive neuroscience (Eisenberg et al., 2019; Enkavi et al., 2019), and (b) can lay the foundation for future work in interpretable models, and cognitive theories (see e.g., Peterson, Bourgin, Agrawal, Reichman, and Griffiths (2021); Miller, Botvinick, and Brody (2021) for recent successes of the transition from predictive to explanatory models).

Related Work

Inhibitory control & test-retest reliability: Behavioral tasks are routinely used to characterize individual differences in putative cognitive processes such as inhibitory control, self-regulation, working memory, and so on. To accurately characterize individual differences in these tasks, we need *reliable* measures – here, reliability describes our ability to rank individuals in a consistent manner across sessions. This is all the more important for longitudinal studies (Garavan et al., 2018). However, multiple systematic studies question the reliability of classically used tasks and their associated indices of individual performance. For instance, Eisenberg et al. (2019) and Enkavi et al. (2019) found that survey measures of self-regulation had higher reliability, and better predicted real-world outcomes, compared to cognitive tasks. Similarly, Hedge et al. (2018) found low test-retest reliability ranging from 0 to 0.82 across seven commonly used inhibition tasks. We believe that the average statistics classically measured in these tasks may not capture the whole of individuals' decision-making policies in the tasks, and instead build predictive models that can also explain some portion of the trial by trial behavioral variation.

Sequential effects in decision-making: Influences of recently experienced stimuli on reaction times have long been known (see e.g., Laming (1968)), and arise even when there is no apparent relevance of stimulus history to behavior on the current trial. Some authors (see e.g., Yu and Cohen (2009)) suggest a model of stimulus statistics estimation under a nonstationary distributional assumption, leading to rational adjustments to the changes in prior expectation. There are competing theories about which particular stimulus history patterns influence behavior, and through what cognitive mechanism (see e.g., Davelaar (2013)). Other work examines the influence of prior stimulus expectation on a within-trial policy tradeoff between various objectives in the task (Leotti & Wager, 2009); however, attempts at defining those within-trial policies vary widely, from descriptive (Ratcliff & Rouder, 1998; Logan & Cowan, 1984b) to prescriptive (Shenoy & Yu, 2011). Some researchers investigate brain regions involved in integrating stimulus-outcome history for policy adjustments (Hwang, Dahlen, Mukundan, & Komiyama, 2017), or suggest entirely alternate causal mechanisms underlying sequential patterns, such as stochasticity in brain processes (Samaha, Iemi, Haegens, & Busch, 2020), or instantaneous coupling between brain areas (Polanía, Moisa, Opitz, Grueschow, & Ruff, 2015).

Computational models of behavior:[3] The drift-diffusion model (DDM (Ratcliff & Rouder, 1998)) is a highly popular mechanistic model of 2-alternative forced choice tasks as well as other similarly structured tasks, and model fits of DDM (e.g., using packages such as HDDM (Wiecki, Sofer, & Frank, 2013)) are often used to extract covariates from behavioral data for correlational analysis with brain signals or clinical outcomes. DDM only models *distributional data* (RT and error distributions), and cannot capture sequential adjustments in behavior; in our experiments, their test-retest reliability is in the ballpark of classical empirical measures.

Closer to our proposal, Dezfouli, Griffiths, Ramos, Dayan, and Balleine (2019) proposed a simple LSTM model for capturing human policy learning in a bandit task, with the goal of classifying subjects into 3 clinically defined groups. In subsequent work (Dezfouli, Ashtiani, et al., 2019), they mapped individual behavioral sequences into a low-dimensional latent space, where, again, the goal was qualitative across-group comparisons of (disentangled) factors underlying decision policies. The embeddings only captured coarse aspects of across-group variation in the task, not sufficient to serve as accurate predictive models of individual behavior.

Chatterjee and Shenoy (2021) proposed learning a feedforward architecture, with subject embeddings, on pooled data from a large number of subjects performing an information seeking task. In their task, while the subject embeddings capture a substantial amount of inter-subject variation in performance, they do not build recurrent models and also do not address the reliability of parameter estimates across sessions.

Our approach: We take a model-agnostic approach to capturing inter-trial variability, using black-box LSTM architectures. Our finding that recurrent architectures capture trial-by-trial responses is evidence that part of the variability is history-driven; further, empirical analyses show that RT and accuracy do in fact tightly covary, again supporting the notion that this captured variability is related to policy adjustments that influence such tradeoffs (Leotti & Wager, 2009). We focus solely on the *predictive task* of capturing RT/choice variation, on a per-subject grain, without limiting learned models to any specific function class. Other models with strong inductive biases (Yu & Cohen, 2009), or aimed at captur-

[3] We only sketch a few approaches relevant to our context: simple cognitive/behavioral tasks, sequential updates to decision policies, and subject-specific parameter estimation.

ing group-level variation (Dezfouli, Ashtiani, et al., 2019) showed poor test-retest reliability in our experiments. Finally, we exploit simple meta-learning architectures to *pool data across subjects* for training – indeed, learning recurrent models from small amounts of single-subject data is difficult in general, and unsuccessful in our setting.

Methods

Eriksen Flanker Task

We describe the Eriksen flanker task used in our experiments (Figure 1). Each trial presents an arrow (left, right) in the middle of the screen, with subjects required to press a button linked with the direction of the arrow, as fast and accurately as possible. Additional, task-irrelevant stimuli are present on either side of the central arrow, in one of 3 trial types – (a) **congruent**: arrows matching the central stimulus, (b) **incongruent**: arrows in opposition to central stimulus, (c) **neutral**: non-arrow stimuli. The expectation is that on congruent trials, the flanking stimuli reinforce the perception of the central stimulus, and on incongruent trials they interfere.

In the task, each subject performs multiple trials per session. For each trial t, with trial type x_t and arrow direction p_t, the subjects' reaction time τ_t and correctness c_t is recorded. We use the notation $\mathbf{x}_t = (x_t, p_t)$ to denote the stimulus, and $\mathbf{r}_t = (c_t, \tau_t)$ to denote the user response for trial t.

Recurrent Networks for Behavior Prediction

Given a sequence of trials $X = \{\mathbf{x}_1 \dots \mathbf{x}_t\}$ experienced by a subject s, we wish to model behavioral responses by training models that mimic behavior, i.e., the model's goal is to predict an individual subject's response sequence $R = \{\mathbf{r}_1 \dots \mathbf{r}_t\}$, rather than maximize the accuracy at the task itself.

As discussed earlier, our goal is to capture sequential adjustments in behavior; we therefore model the response at time t as a function of all previous stimuli and feedback. To model this temporal dependency, we use an LSTM (Hochreiter & Schmidhuber, 1997) to integrate stimulus and response history over the past trials.

Our proposed model RECNET, consisting of an LSTM accumulator and a feedforward response prediction network, is shown in Figure 2a. The LSTM has as its inputs, the previous stimulus $\mathbf{x}_{t-1}$ and the previous responses $\mathbf{r}_{t-1}$. We combine the LSTM's hidden state $\mathbf{h}_t$ and the current trial's stimulus information $\mathbf{x}_t$ using a feedforward network, and predict the user's choice ($\widehat{c}_t$) and RT ($\widehat{\tau}_t$) on the current trial.

The model parameters Θ are learned by minimizing the following objective:

$$L\left(\widehat{R}; X, \Theta\right) = \sum_{i=1}^{t} l_{CE}\left(\widehat{c}_t, c_t\right) + l_{Huber}\left(\widehat{\tau}_t, \tau_t\right) \qquad (1)$$

Here, l_{CE} is the cross-entropy loss and l_{Huber} is the Huber loss. $\widehat{R}$ is defined in the same way as R but for the model predictions. The true user responses R, also an input to the loss function, are elided here and in the following sections for notational simplicity.

Note that we use the subject's actual responses $\mathbf{r}_{t-1}$ as input to the LSTM, rather than the model's predictions on the previous time step. This is because we are specifically interested in capturing the *policy adjustment* as a function of sequentially integrated history; using the model's own predictions may result in accumulative noise over the sequence of predictions for the subject, and hinder learning.

Figure 2b shows a simplification of the model by eliminating the recurrent component (LSTM)–we use this purely feedforward model, FFNET, as a baseline for comparison. FFNET is also trained using the loss function described above. Note that since the model has no contextual information distinguishing any given trial from any other trial with the same stimuli $\mathbf{x}_t$, it can only learn the average error rates and response times broken down by trial types. The primary role of FFNET as a baseline is to account for all variance in choice explained by trial type alone, i.e., without the use of sequential adjustments. Figure 2c considers an intermediate model; one that uses a flat prediction structure like FFNET but also takes as input the immediately preceding time point's stimuli and responses. This model is named FFPREV and is used to re-examine previous findings that simple linear models (e.g., logistic regression) show sequential influence of the immediately preceding time point, but typically not further back in the past (Hwang et al., 2017).

Meta-Learning via Parametrization

We propose a simple meta-learning approach (e.g., Figure 2d), wherein the network parameters are shared across subjects, and each subject s is parametrized using a subject-specific embedding $\theta_{SE}^{(s)}$ passed as additional input to the model. These *learned embeddings* allow for the meta-model to be customized, to some degree, to individual subjects. We refer to the model in Figure 2d as RECNET(SE) and train it using data pooled across all subjects, i.e.,

$$\mathcal{L}_{SE} = \sum_{s} L\left(\widehat{R}^{(s)}; X^{(s)}, \Theta, \Theta_{SE}\right) \qquad (2)$$

where $X^{(s)}, \widehat{R}^{(s)}$ refer to the trial inputs and responses from subject s, respectively, and Θ_{SE} is the set of subject-specific embeddings $\theta_{SE}^{(s)}$ for all subjects.

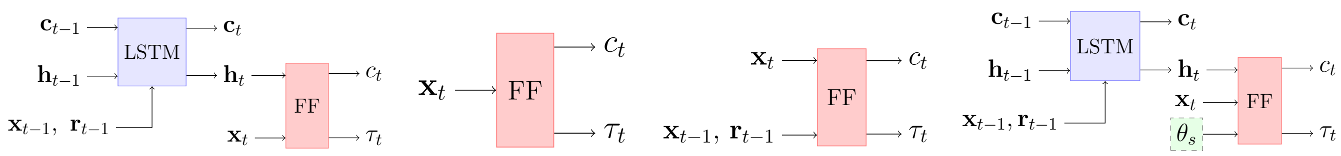
We note that the other baseline models FFNET and FFPREV can also be similarly enhanced via pooled training and subject parameters, creating the models FFNET(SE) and FFPREV(SE) respectively.

Experiments

Dataset: We reanalyze data collected by Hedge et al. (2018)[4], which includes two sessions about 3 weeks apart. Each subject completed 720 trials (240 in each condition - congruent, incongruent, neutral) per session. The study included two different batches, of 50 and 62 subjects each. We pool data from both giving 107 subjects[5]. For each subject

[4] released under CC BY 4.0, `https://osf.io/cwzds/`

[5] The remaining 5 did not return for the second session.



(a) RecNet (b) FFNet (c) FFPrev (d) RecNet(SE)

Figure 2: Recurrent models of decision policy adjustment. (a) A recurrent model RecNet, with an LSTM to integrate trial & response history (previous trial stimuli $\mathbf{x}_{t-1}$ and response $\mathbf{r}_{t-1}$), and a feedforward network (FF) to combine with current trial stimuli $\mathbf{x}_t$ and predict current choice (c_t) & RT (τ_t). ($\mathbf{c}_t$ and $\mathbf{h}_t$ denote the cell state and the hidden state of the LSTM.) (b) Non-recurrent baseline for comparison–FFNet–a purely feedforward network with current trial inputs. (c) Non-recurrent baseline with bounded history–FFPrev, similar to FFNet, except that it takes directly as input the previous trial's information ($\mathbf{x}_{t-1}, \mathbf{r}_{t-1}$). (d) The meta-model RecNet(SE) which can be trained over pooled data from multiple subjects, and includes a learned subject embedding (SE) for customization to individual subject data. Other models can be similarly customized.

& session, the first 80% of the trials are used for training and the remaining for validation *i.e.*, two sets of 576 & 144 trials, which we label `S1Train`, `S1Val`, `S2Train`, and `S2Val`.

Training & validation: All models are first trained using `S1Train` and evaluated on `S1Val`, allowing us to compare their predictive power. For models that have subject parameters, we evaluate transfer using the following process: the shared parameters (network weights) are kept frozen from the first session's training, and only subject parameters are re-estimated using `S2Train` data. The frozen meta-models, plus the re-estimated subject parameters, are then evaluated on `S2Val`. Finally, the two sets of subject parameters are compared for estimating across-session reliability.

Architecture and Training Details: A single layer LSTM with an input size of 6 and hidden dimension 15 is used. The feedforward network consists of two parallel sets of layers, one each for predicting the choice and RT. These are two-layer fully connected networks having 15 hidden units and ReLU activation. Training is done using Adam optimizer (Kingma & Ba, 2015) using a learning rate of 0.003 for session 1 and 0.001 for session 2, for a maximum of 2000 epochs, with early stopping. A weight decay (L2 regularization) of 0.001 is used for the subject embeddings (Θ_{SE}). A batch size of 2 is used. We use log(RT) instead of RT since the log distribution is easier to learn.

Metrics and evaluation: We calculate within-subject correlations between predicted and actual RT, and predicted choice (likelihood) and actual choice. These correlations are averaged across subjects; we also keep track of the number of subjects for which the correlation was statistically significant (setting $p < 0.05$). We compare two models using paired t-tests on the subject-grain correlation coefficients, and assess statistical significance. All reported metrics are on the validation splits alone. We also compared models on MSE (RT predictions) and cross-entropy (choice); the findings were broadly similar and we do not report those metrics here. For simplicity, we evaluate models without subject parameters only on session 1 data; results are similar on session 2 data.

Test-retest reliability: We compute the intra-class correlations (ICC (Shrout & Fleiss, 1979)) of estimated subject parameters across the two sessions. ICCs are used as a measure of consistency or reproducibility of measurements made by different raters for the same subject. It roughly compares the variation of ratings of the same subject to the total variation across all ratings and all subjects.

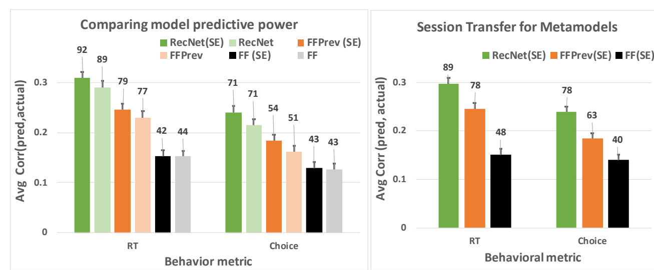


Figure 3: Model predictive power on session 1 data (left panel), and transfer performance on session 2 data (right panel, see text for details). RecNet(SE) significantly outperforms ($p <$1e-10, pairwise comparisons), and each model class is improved by adding subject parameters. For meta-models, simply re-estimating embeddings on session 2's data provides performance equivalent to that of session 1.

Results

Sequential policy updates: Figure 1b relates response times with accuracy on congruent & incongruent trials in the task. We divided each subject's trials into decile buckets by RT, and pooled data across subjects to calculate accuracy by decile bucket. The data shows a clear tradeoff between response speed and accuracy, showing that RT variations are not merely perceptual or motor noise, but reflect adjustments to the decision policy.

Model fits to behavior: Figure 3(a) shows the averaged per-subject correlation between model predictions and actual data, for RT (left group) and choice (right group). Also shown are error bars (SEM across subjects) and the number of subjects with statistically significant correlations ($p < 0.05$) (numbers on top of the bars). The models listed include baselines (FFNet, FFPrev, RecNet) trained on pooled

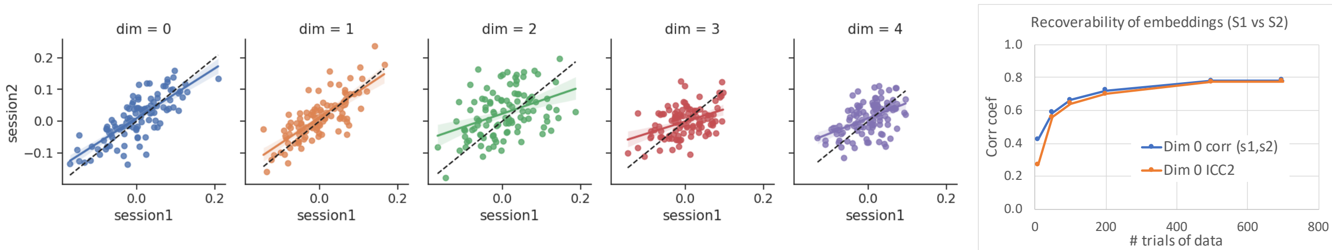


Figure 4: Recoverability of subject parameters for RECNET(SE). Each point represents the corresponding parameter estimate for one subject, compared across sessions. The meta-model and subject parameters were learned on session 1, and subject parameters alone re-estimated using data from session 2. The first 5 panels show the 5 dimensions of the embedding. Panel 6 shows the correlation between sessions for dimension 0 alone, as a function of #trials of session 2 used for estimation.

data across subjects, and their corresponding meta-models (FFNET(SE), FFPREV(SE), RECNET(SE), which include subject parameters for personalized predictions). The following trends are clearly demonstrated: **(1)** No-history models perform poorly; FFNET and FFNET(SE) both show low average correlations, and are similar to each other, **(2)** Even a single previous timestep provides substantial information (FFPREV, FFPREV(SE))–this is consistent with previous findings (Hwang et al., 2017), **(3)** Recurrent models clearly outperform the other models (RECNET and RECNET(SE)). Finally, each model family's performance is consistently improved by the addition of subject parameters.

We compared the meta-models in pairwise fashion using paired t-tests: RECNET(SE), FFPREV(SE), FFNET(SE); each model was statistically significantly better than the next with high confidence ($p < 1e-10$ for all comparisons).

Generalization: To test generalization to a second session, we kept the meta-models fixed, and only re-estimated subject parameters on a portion of the new session's data, starting with random initialization. Figure 3(b) shows that each model's performance on the new session's data is nearly identical to performance on the original session's validation data; this shows excellent transference of predictive power across sessions. This is a significant result: *given a pre-learned meta-model, say* RECNET(SE)*,* we only need to estimate a small number of subject parameters per subject to be able to predict behavior on an entirely different session of behavioral data. This also means that parameters for entirely new subjects can be easily estimated using a prelearned meta-model.

Repeatability: Figure 4 compares the estimated subject parameters (here, of dimension 5) for RECNET(SE) across sessions; each plot compares one dimension in isolation across subjects. Parameter estimates are strongly correlated, particularly dimensions $\{0,1\}$. The rightmost panel of Figure 4 shows the correlation across sessions for dimension 0, as a function of # trials from session 2 used for estimating subject parameters. A small amount of data (around 200 trials) is already sufficient to obtain a stable re-estimate.

Intraclass Correlations: Table 1 presents, for each subject-specific index of behavior, its ICC (a measure of across-session reliability–see Methods), the associated confidence interval, and the Pearson correlation coefficient (closely related to ICC). For model parameters, we only present the two dimensions that had the highest ICC scores; these are also typically (but not necessarily) the first 2 dimensions. We also evaluated classical measures in the task, based on RT and accuracy by trial type. RT cost and Accuracy cost refer to the difference in averages (RT, accuracy respectively) between congruent and incongruent trials.

As seen in Table 1, ICC scores are ordered similar to model predictive power; in particular, RECNET(SE) has clearly superior test-retest reliability, driven by its capturing of sequential policy adjustments as a subject-specific trait. This also indirectly supports our primary motivation–that strongly predictive models correspond to more reliable subject indices. FFPREV(SE) shows intermediate ICC scores, since it has bounded access to history. Interestingly, FFNET(SE) closely matches ICCs of empirical/behavioral measures in the task, since it is only able to capture average performance by trial type, similar to the empirical measures.

Finally, referring back to Figure 4 (last panel), we see that ICC for dimension 0 already exceeds all behavioral measures using as few as 200 trials; this is especially important given the claim that estimation noise may be a major factor underlying poor retest reliability of classical measures, requiring large numbers of trials (Rouder, Kumar, & Haaf, 2019).

Other computational models: Due to paucity of space, we only summarize key findings from the study of other computational models. We evaluated the fitted parameters of the Drift-Diffusion Model (DDM, a popular mechanistic model for simple decision-making tasks), using the HDDM package (Wiecki et al., 2013) and found them to be on par with behavioral measures ($\sim$0.59). This is unsurprising, as DDMs only model the overall distribution (i.e., average statistics) and cannot capture across-trial variance.

We similarly evaluated the hypernetwork proposed by Dezfouli, Ashtiani, et al. (2019) and found the latent representations to have very low ICCs ($\sim$0.23). This approach, although conceptually similar in the use of a recurrent network, pooled learning, and latent parameters, was designed primarily for qualitative separation (disentangled representations) and across-group comparisons, rather than accurate subject-specific predictive modeling. Finally, we evaluated a popular proposal for modeling sequential RT variations–the Dynamic

Bayes Model (Yu & Cohen, 2009)–and found its estimated parameters to also have low reliability. Here, too, there is a mismatch of modelers' intent and application, since they primarily addressed the nonstationary estimation of stimulus frequencies using a specific inductive bias, whereas we use models with sufficient representational capacity and no inductive constraints (LSTMs) with the primary objective of predictive accuracy.

Examining learned model parameters: We display the cross-correlation matrix between all empirical measures and all embedding dimensions for RECNET(SE) in Figure 5,. The dimensions SE0 and SE1 for RECNET(SE) have the highest ICC; they appear to show moderate-to-high positive (resp. negative) correlation across all behavioral measures. Notably, they show similar correlation patterns as the incongruent accuracy, which is the most reliable of the empirical measures; this is despite the fact that RECNET(SE) models sequential adjustments, a source of information not (directly) available to the average empirical metrics. This suggests that average metrics contain some coarse representation or correlates of the decision policies implicit in RECNET(SE); further analysis is needed to explicate these relationships.

Table 1: Intraclass-correlations, confidence intervals, and across-session correlations for classical task measures and learned subject embeddings. Primary classical metrics (RT/Accuracy Cost) have low ICCs; incongruent accuracy is better with ICC 0.70. RecNet(SE) embeddings show clearly higher ICCs, and are thus more reliable as markers of individual behavior. All correlations except RT cost are $p <$1e-10.)

		ICC	CI 95%	Corr
RECNET(SE)	SE:top1	**0.77**	[0.68, 0.84]	**0.78**
	SE:top2	0.75	[0.65, 0.83]	0.77
FFPREV(SE)	SE:top1	0.71	[0.59, 0.80]	0.73
	SE:top2	0.62	[0.48, 0.72]	0.62
FFNET(SE)	SE:top1	0.62	[0.48, 0.72]	0.65
	SE:top2	0.62	[0.46, 0.73]	0.66
Classical RT	congruent RT	0.62	[0.46, 0.73]	0.66
	neutral RT	0.58	[0.41, 0.71]	0.64
	incongruent RT	0.57	[0.38, 0.70]	0.63
	RT Cost	0.23	[0.05, 0.40]	0.27
Classical Acc	congruent Acc	0.53	[0.37, 0.66]	0.59
	neutral Acc	0.52	[0.34, 0.65]	0.58
	incongruent Acc	0.70	[0.57, 0.80]	0.74
	Accuracy Cost	0.57	[0.42, 0.68]	0.56

Discussion

We outlined the problem of reliable session-to-session measurement of individual behavior in inhibitory control tasks. We showed the promise of modeling sequential policy adjustments in the task, a direction not explored in classical analyses of task behavior for identifying subject-specific indices of performance. We proposed meta-learning of recurrent models for capturing sequential choice and response time behavior, and showed that these models were significantly more predictive than per-subject models, or non-recurrent models with bounded history. Our approach provides, along with a meta-model for predicting behavior in the task, subject parameters that serve as compact, individual parametrizations of the meta-model to subject-specific data. The learned subject parameters have high test-retest reliability and can be parsimoniously estimated with only a few trials of data for a new subject. This is especially important given findings in the literature that behavioral estimates are noisy and may need significant amounts of data for estimation (Rouder et al., 2019). Our subject parameters show promise as a richer window into individual differences in inhibitory control, and may better capture across-group differences for diagnosing disruptions in decision-making faculties.

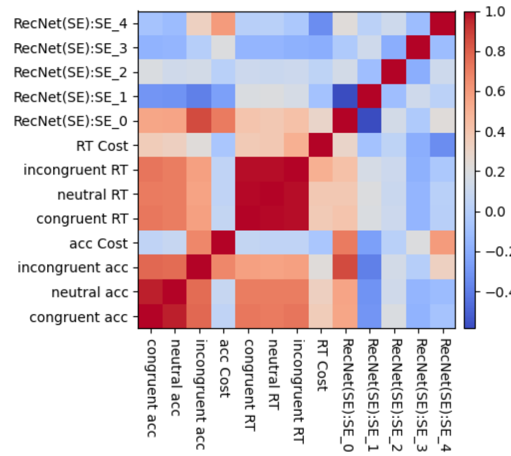


Figure 5: Cross-correlations between classical task metrics and learned subject parametrizations.

From the perspective of developing theories of cognition and decision-making, there are two straightforward next steps for our work. On one hand, the meta-models must be capturing a *parametrized family of behavioral policies*, using some functional form relevant to ecological and statistical concerns. In general, interpreting and understanding the functions learned by black-box models such as ours is a very promising direction of study given their strong predictive power. On the other hand, many recent papers have attempted nonparametric analyses to identify potential *latent dimensions* or factors that describe unitary constructs of inhibitory control that generalize across many such diagnostic tasks; see for example (Eisenberg et al., 2019; Gärtner & Strobel, 2021).

We can easily extend our models to a multi-task setting, and study the induced subject parameters that span these tasks for clues about the structure of inhibitory control. We believe that by using predictive accuracy as a strong supervisory signal, we may uncover much more informative latent spaces, and consequently, a better understanding of the conceptual construct of "inhibitory control" in general.

References

Chatterjee, S., & Shenoy, P. (2021, May). Model-agnostic fits for understanding information seeking patterns in humans. *Proceedings of the AAAI Conference on Artificial Intelligence*, *35*(1), 784-791.

Davelaar, E. J. (2013). When the ignored gets bound: Sequential effects in the flanker task. *Frontiers in psychology*, *3*, 552.

Dezfouli, A., Ashtiani, H., Ghattas, O., Nock, R., Dayan, P., & Ong, C. S. (2019). *Disentangled behavioral representations* (Tech. Rep.).

Dezfouli, A., Griffiths, K., Ramos, F., Dayan, P., & Balleine, B. W. (2019, June). Models that learn how humans learn: The case of decision-making and its disorders. *PLoS Comput. Biol.*, *15*(6), e1006903.

Donders, F. C. (1969). On the speed of mental processes. *Acta Psychol.*, *30*, 412.

Eisenberg, I. W., Bissett, P. G., Enkavi, A. Z., Li, J., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Uncovering the structure of self-regulation through data-driven ontology discovery. *Nature communications*, *10*(1), 1–13.

Enkavi, A. Z., Eisenberg, I. W., Bissett, P. G., Mazza, G. L., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Large-scale analysis of test–retest reliabilities of self-regulation measures. *Proceedings of the National Academy of Sciences*, *116*(12), 5472–5477.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & psychophysics*, *16*(1), 143–149.

Garavan, H., Bartsch, H., Conway, K., Decastro, A., Goldstein, R., Heeringa, S., … Zahs, D. (2018). Recruiting the abcd sample: Design considerations and procedures. *Developmental cognitive neuroscience*, *32*, 16–22.

Gärtner, A., & Strobel, A. (2021). Individual differences in inhibitory control: A latent variable analysis. *Journal of Cognition*, *4*(1).

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior research methods*, *50*(3), 1166–1186.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735–1780.

Hwang, E. J., Dahlen, J. E., Mukundan, M., & Komiyama, T. (2017). History-based action selection bias in posterior parietal cortex. *Nature communications*, *8*(1), 1–14.

Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. In *3rd international conference on learning representations, ICLR 2015, san diego, ca, usa, may 7-9, 2015, conference track proceedings.*

Laming, D. R. J. (1968). Information theory of choice-reaction times.

Lappin, J. S., & Eriksen, C. W. (1966). Use of a delayed signal to stop a visual reaction-time response. *Journal of Experimental Psychology*, *72*(6), 805.

Leotti, L. A., & Wager, T. D. (2009). Motivational influences on response inhibition measures. *J. Exp. Psychol. Hum. Percept. Perform.*.

Logan, G. D., & Cowan, W. B. (1984a). On the ability to inhibit thought and action: A theory of an act of control. *Psychological review*, *91*(3), 295.

Logan, G. D., & Cowan, W. B. (1984b). On the ability to inhibit thought and action: A theory of an act of control. *Psychol. Rev.*, *91*(3), 295.

Miller, K. J., Botvinick, M. M., & Brody, C. D. (2021). From predictive models to cognitive models: Separable behavioral processes underlying reward learning in the rat. *bioRxiv*, 461129.

Peterson, J. C., Bourgin, D. D., Agrawal, M., Reichman, D., & Griffiths, T. L. (2021). Using large-scale experiments and machine learning to discover theories of human decision-making. *Science*, *372*(6547), 1209–1214.

Polanía, R., Moisa, M., Opitz, A., Grueschow, M., & Ruff, C. C. (2015, August). The precision of value-based choices depends causally on fronto-parietal phase coupling. *Nat. Commun.*, *6*, 8090.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychol. Sci.*, *9*, 347–356.

Rouder, J., Kumar, A., & Haaf, J. M. (2019). Why most studies of individual differences with inhibition tasks are bound to fail.

Samaha, J., Iemi, L., Haegens, S., & Busch, N. A. (2020, August). Spontaneous brain oscillations and perceptual Decision-Making. *Trends Cogn. Sci.*, *24*(8), 639–653.

Shenoy, P., & Yu, A. J. (2011). Rational decision-making in inhibitory control. *Front. Hum. Neurosci.*, *5*(48).

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin*, *86*(2), 420.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology*, *18*(6), 643.

Victor, T. A., Khalsa, S. S., Simmons, W. K., Feinstein, J. S., Savitz, J., Aupperle, R. L., … Paulus, M. P. (2018). Tulsa 1000: a naturalistic study protocol for multilevel assessment and outcome prediction in a large psychiatric sample. *BMJ open*, *8*(1), e016620.

Wiecki, T., Sofer, I., & Frank, M. (2013). Hddm: Hierarchical bayesian estimation of the drift-diffusion model in python. *Frontiers in Neuroinformatics*, *7*, 14. doi: 10.3389/fninf.2013.00014

Yu, A. J., & Cohen, J. D. (2009). Sequential effects: Superstition or rational behavior? *Adv. Neural Inf. Process. Syst.*, *21*, 1873–1880.