



Recovering Reliable Idiographic Biological Parameters from Noisy Behavioral Data: the Case of Basal Ganglia Indices in the Probabilistic Selection Task

Yinan Xu¹ · Andrea Stocco²

Accepted: 25 February 2021 / Published online: 24 March 2021
© Society for Mathematical Psychology 2021

Abstract

Behavioral data, despite being a common index of cognitive activity, is under scrutiny for having poor reliability as a result of noise or lacking replications of reliable effects. Here, we argue that cognitive modeling can be used to enhance the test-retest reliability of the behavioral measures by recovering individual-level parameters from behavioral data. We tested this empirically with the Probabilistic Stimulus Selection (PSS) task, which is used to measure a participant's sensitivity to positive or negative reinforcement. An analysis of 400,000 simulations from an Adaptive Control of Thought-Rational (ACT-R) model of this task showed that the poor reliability of the task is due to the instability of the end-estimates: because of the way the task works, the same participants might sometimes end up having apparently opposite scores. To recover the underlying interpretable parameters and enhance reliability, we used a Bayesian Maximum A Posteriori (MAP) procedure. We were able to obtain reliable parameters across sessions (intraclass correlation coefficient ≈ 0.5). A follow-up study on a modified version of the task also found the same pattern of results, with very poor test-retest reliability in behavior but moderate reliability in recovered parameters (intraclass correlation coefficient ≈ 0.4). Collectively, these results imply that this approach can further be used to provide superior measures in terms of reliability, and bring greater insights into individual differences.

Keywords Probabilistic Stimulus Selection task; · Reliability test; · Basal ganglia; · Direct and indirect pathways; · Computational modeling; · ACT-R

Introduction

To understand cognition, it is important that the behavioral measures that we use to indirectly index brain function are valid and reliable. Unfortunately, this is often not the case, with published effects often showing low replicability or task results having poor reliability across time (Open Science Collaboration, 2015). As noted by Hedge et al. (2018), low reliability might sometimes be due to the use of tasks (such as

Flanker, Stroop, stop-signal, and go/no-go) that are explicitly designed to have robust group effects by low variability. Idiographic (i.e., individual-level) parameters in cognitive modeling, on the other hand, can capture individual-level characteristics (Daw, 2011; Collins, 2018) and have high test-retest reliability. For instance, there is substantial evidence that drift-diffusion models fitted to data from individual participants can disentangle the effects of the different processes driving behavior, and enhance our understanding on how neural activity varies across individuals (White et al., 2014, 2016). Furthermore, Sense et al. (2016, 2018) have repeatedly shown that long-term memory rate of forgetting, estimated through a modified exponential decay model, is stable across sessions and across materials.

In this paper, we argue that idiographic parameters in cognitive modeling can be used to enhance the reliability of behavioral measures, and specifically of behavioral measures that are supposed to track underlying neurobiological characteristics. Specifically, we show that cognitive models can be used to reliably recover the values of underlying

✉ Yinan Xu
yinanx97@gmail.com

Andrea Stocco
stocco@uw.edu

¹ University of Houston, Houston, TX, USA

² Department of Psychology & Institute for Learning and Brain Sciences, University of Washington, Seattle, WA, USA

parameters (which reflect neurocognitive processes) even when the behavioral data itself is noisy and lacks replicability. As an example, we will use the Probabilistic Stimulus Selection task (PSS), an experimental paradigm that has been widely adopted in neuroscience research to investigate basal ganglia function (Frank et al., 2004) but whose effectiveness has recently come under scrutiny (Baker et al., 2013; Grogan et al., 2017). Schutte et al. (2017) also raised the possibility that the task suffers from a perceptual discriminability problem. In other words, some task stimuli (Fig. 1) appear to be more clearly discriminable than others, leading to a perceptual preference issue which causes participants to choose in a biased way. Thus, differences in perceptual discriminability between individual stimuli strongly influenced behavior results, hence causing an unstable reliability.

To this end, reliability tests on behavioral data and model parameters were run consecutively on the original PSS task. A modified version of the task was introduced as a follow-up experiment with reliability tests on both behavioral data and model recovered parameters. In both studies, we will show that the use of cognitive modeling can (a) shed light on the nature of discrepant findings by different laboratories and (b) recover interpretable, idiographic parameters from otherwise noisy behavioral data, providing superior measures of validity and reliability and greater insight into individual differences.

The Probabilistic Stimulus Selection (PSS) Task

The task examined herein is the Probabilistic Stimulus Selection (PSS) task. The PSS task is an iterative, forced-choice, implicit decision-making paradigm first introduced by Frank et al. (2004) in which participants are asked to repeatedly choose from pairs of non-verbalizable stimuli, each of which has a different probability of giving a reward (varying linearly from 20 to 80%). The task consists of a training phase and a testing phase. During the training

phase, the participants are trained to select the most rewarding stimulus out of three different fixed pairs (Fig. 1, left). The highest probability stimulus is always paired with the lowest one, then the second-highest stimulus with the second-lowest one, and the third-highest probability stimulus pairs with the third-lowest one. Feedback about the outcome of their selection (that is, whether it resulted in being rewarded or not) is shown on the screen immediately after each selection. To discourage the participants from using explicit strategies (for example by keeping a running total of each stimulus’s history of successes), the stimuli are intentionally designed to be difficult to verbalize and memorize: they are represented as Hiragana characters from the Japanese writing system and are presented solely to non-Japanese speakers (for simplicity, the stimuli will be indicated with the letters *A, B ... F*, as in Fig. 1). The learning occurring in this training phase is then assessed in the testing phase, where the six stimuli are now combined into all possible pairs (Fig. 1, right) and feedback is not given to prevent further learning.

Note that, during the training phase, participants might learn equally well by either learning to choose the most rewarding stimuli (i.e., *A*) or by avoiding the least rewarding ones (i.e., *B*). These two processes can be distinguished in the testing phase by calculating two measures, Choose and Avoid accuracies. Choose accuracy is calculated as the probability of choosing *A* while paired with *C, D, E, or F*; Avoid accuracy, on the other hand, is the probability of choosing *C, D, E, and F* over *B*. Thus, Choose and Avoid accuracies are calculated from non-overlapping stimuli pairs (*A-C, A-D, A-E, A-F* vs. *B-C, B-D, B-E, B-F*) that were not experienced in the learning phase and do not include the direct contrast pair *A-B*. Therefore, they are operationally independent.

The importance of this task lies in the fact that Choose and Avoid accuracies provide insight into a person’s biology, and, specifically, into the physiology of the basal ganglia. The basal ganglia are a set of subcortical nuclei that modulate the

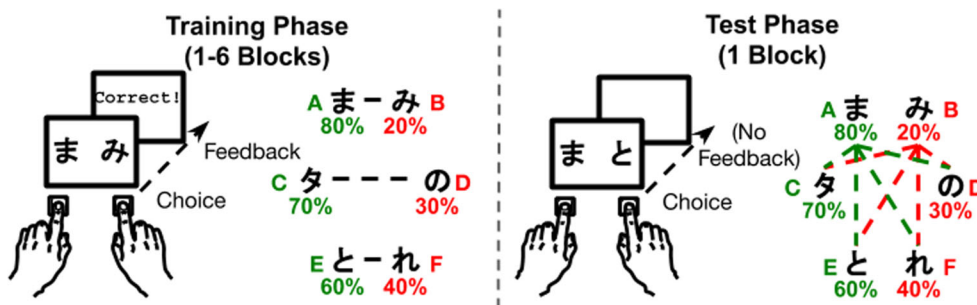


Fig. 1 Overview of the Probabilistic Stimulus Selection task. During the training phase (left), subjects are asked to repeatedly select one stimulus from the three possible pairs. The feedback received (“Correct!” or “Incorrect!”) depends on the stimulus chosen and is shown immediately after each choice. The six stimuli are presented in fixed pairings. During

the testing phase (right), subjects perform the same task as the training phase but without the feedback. The stimuli now appear in new pairings that include either the most rewarding stimulus (green lines) or the least rewarding stimulus (red lines) against each of the remaining stimuli

activity of the prefrontal cortex and are involved in many cognitive functions, most importantly in acquiring procedural knowledge (Knowlton et al., 1996; Yin & Knowlton, 2006). The connections between these nuclei are organized into two pathways, called the direct and the indirect pathway, which have opposite effects on cortical activity (Albin et al., 1989; DeLong, 1990). While the direct pathway exerts an excitatory effect on the prefrontal cortex, the indirect pathway has an inhibitory influence. The striatal neurons that originate the two pathways also express different dopamine receptors: while the direct pathway neurons express D1 receptors that are excited by dopamine release, indirect pathway neurons express D2 receptors and are inhibited by dopamine (Gerfen et al., 1990). Because dopamine is important in reward-based learning and decision-making, it was hypothesized that Choose accuracy reflects the contribution of the direct pathway and Avoid accuracy reflects the activity on the indirect pathway. In fact, the original study by Frank et al. (2004) shows that, when Parkinson's Disease (PD) patients are on dopamine-promoting medication, they are more likely to be "Choosers," meaning their Choose accuracy is higher than the Avoid accuracy. Correspondingly, patients are more likely to be "Avoiders" when they are off-medication and their dopamine level is low. This pattern of results is broadly consistent with a great number of known studies suggesting dopaminergic drugs selectively modulate learning from positive outcomes while impairing negative learning (Rutledge et al., 2009; Mathar et al., 2017; Pessiglione et al., 2006; Shohamy et al., 2008). Furthermore, while boosting dopamine transmission in Parkinson's Disease (PD) patients improved reward learning but worsened punishment avoidance, blocking dopamine transmission in Tourette's Syndrome (TS) patients favored punishment avoidance but impaired reward seeking, extending previous findings in PD to another pathological condition, TS (Palminteri et al., 2009). Voon et al. (2010) also stated that dopaminergic medications again impaired learning from negative outcomes but increased learning from positive ones in PD, and this was related to an increase of striatal prediction error activity, hence a "better than expected" result, suggesting a learning mechanism for the test effects. Same result was also found for non-ICD patients (Djamshidian et al., 2010; Cools et al., 2006; Piray et al., 2014).

Further evidence has bolstered the original claims of an association between Choose and Avoid accuracies and basal ganglia function. The original pattern of results in PD patients was replicated by Frank et al. (2007a). Frank et al. (2007b), for example, showed that higher Choose accuracy is associated DARPP-32 gene polymorphisms that promote the expression of D1 dopamine receptors on the direct pathway, and higher Avoid accuracy is associated DRD2 gene polymorphisms that promote expression of D2 dopamine receptors on the indirect pathway. Waltz et al. (2007) and Cicero et al. (2014) showed that that patients affected by schizophrenia, a neuropsychiatric

disorder associated with low striatal dopamine, also exhibit lower Choose accuracies than controls, much like unmedicated PD patients. Similarly, Endrass et al. (2011) found that patients suffering from obsessive-compulsive disorder exhibit higher Avoid accuracies than controls, which is also consistent with low tonic dopamine. Finally, and consistent with the decrease of dopamine in ageing, younger adults exhibited a stronger bias towards Choose than older adults (Simon et al., 2010).

Thus, because of its ability to connect relatively inexpensive behavioral measures with a meaningful biological substrate, the PSS has become a prominent task in experimental brain research. However, the PSS task's reliability has recently been called into question. For example, an experiment by Grogan et al. (2017) with PD patients failed to reproduce the original effects of dopaminergic medications on PSS performance. Also, Baker et al. (2013) found no evidence that patterns of behavior are stable in this task over time. In their study, they conducted a test-retest reliability analysis on the PSS task performance on 90 undergraduate students. This result showed poor reliability of the behavior measures in indexing cognitive processes in reinforcement learning, with virtually no correlation between an individual's tendency to be either an Avoider or Chooser across consecutive sessions.

Summary

In summary, although existing literature suggests that the PSS task can successfully track the function of the basal ganglia's direct and indirect pathways, and the task has been therefore vastly used for this purpose, the reliability of the task needs to be further determined. To deal with this matter, we conducted a new reliability test (Experiment 1) using the same versions of the PSS task used in the original Frank et al.'s (2004) study, and analyzed the behavioral data to measure internal and test-retest reliability. We also proceeded to further examine the factors affecting task reliability by using a biologically-plausible computational model to simulate performance on the PSS task across a range of realistic biological parameters. Furthermore, we combined both approaches, and investigated whether the predicted model performance could be used to recover important individual differences information that would be superior, in terms of reliability, to the pure behavioral measure. Specifically, a Maximum A Posteriori (MAP) procedure was done to recover the underlying biological parameters of the model from the behavioral data. Finally, to address the limitation of generalizability of the original task, and to ensure our model-guided approach is reliable on recovering individual differences, we repeated the above procedures in a second study (Experiment 2) in which the original Japanese Hiragana characters were replaced with different, non-verbalizable typefaces.

Experiment 1

Methods and Materials

Participants Seventy-seven healthy participants (age 18–30, 41 females) from the University of Washington’s undergraduate population took part in the experiment in exchange for credit course. All participants completed two sessions of the PSS exactly 1 week apart; with the second session always occurring on the same day of week and at the same time of day as the first session. All of the study procedures were approved by the University of Washington Institutional Review Board (Application #00000820). Data from six participants was discarded because they were familiar with Japanese, and could thus verbalize the Hiragana stimuli.

Task All participants completed the PSS task, using the same software used by Frank et al. (2004). At the beginning of the session, participants were asked to place their left index finger on the key corresponding to the digit “1” and their right index finger on the key corresponding to the digit “0” of a standard computer keyboard placed in front of them. During the various phases of the task, pairs of Hiragana characters were shown on the screen (Fig. 1). For each trial, participants were required to choose one of the two characters by pressing the key (“1” or “0”) corresponding to the character’s position on the screen (left or right). Participants were explicitly told that, within each pair, one character was more likely to be “correct” than the other, although probabilistically so, and invited to follow their instinct in choosing which one. After they had pressed a button, the two characters disappeared. During the training phase, feedback was immediately presented on the screen as “Correct!” in blue, or “Incorrect!” in red. If the participant didn’t press any button within 6 seconds, “no response detected” in red color was shown on the screen. This was to ensure that the subject was engaging in the task and also to discourage the subject from using explicit methods to remember the patterns rather than learning them through trial and error. After a 1-s crosshair, a new pair of characters was presented.

During the training phase, trials were presented in blocks of 60 each (20 for each of the three training pairs: Fig. 1), with a self-paced pause in-between. At the end of each block, each participant’s performance was re-assessed and, if it passed a predefined learning criterion, participants moved on to the testing phase. If participants had not passed the learning criterion by the end of the sixth block, they would move on to the testing phase anyway. To satisfy the learning criterion, participants had to select the best option within each pair with at least the following probabilities: 50% of *E* in the *E-F* pairs, 60% of *C* in the *C-D* pairs, and 65% of *A* in the *A-B* pairs.

During the testing phase, a single block of 60 trials was presented, with no feedback after each choice (Fig. 1). The 60

trials consisted of four presentations of each of the 15 possible pairs of the six stimuli. Choose and Avoid accuracies were calculated from two non-overlapping subsets of these trials.

Results

Split-Test Reliability First, we examined the split-test reliability of these measures. This was done by separately calculating the values of the two main variables, Choose and Avoid, for different pairs of stimuli, depending on whether *A* and *B* were presented on the left (e.g., “*A-C*”, “*B-C*”, etc.) or on the right side of the screen (e.g., “*C-A*”, “*C-B*”, etc.). These measures were called Choose Left, Avoid Left, Choose Right, and Avoid Right, respectively. The Pearson correlation coefficient between the Left and Right version of each measure was then calculated. As shown in Fig. 2, the split-test correlation coefficients of Choose and Avoid were significant in both sessions. Specifically, we found a positive correlation of Choose Left and Choose Right in Session 1 [$r(71) = .44, p < .001$] and in Session 2 [$r(71) = .40, p < .001$] and between Avoid Left and Avoid Right in Session 1 [$r(71) = .46, p < .001$] and in Session 2 [$r(71) = .40, p < .001$].

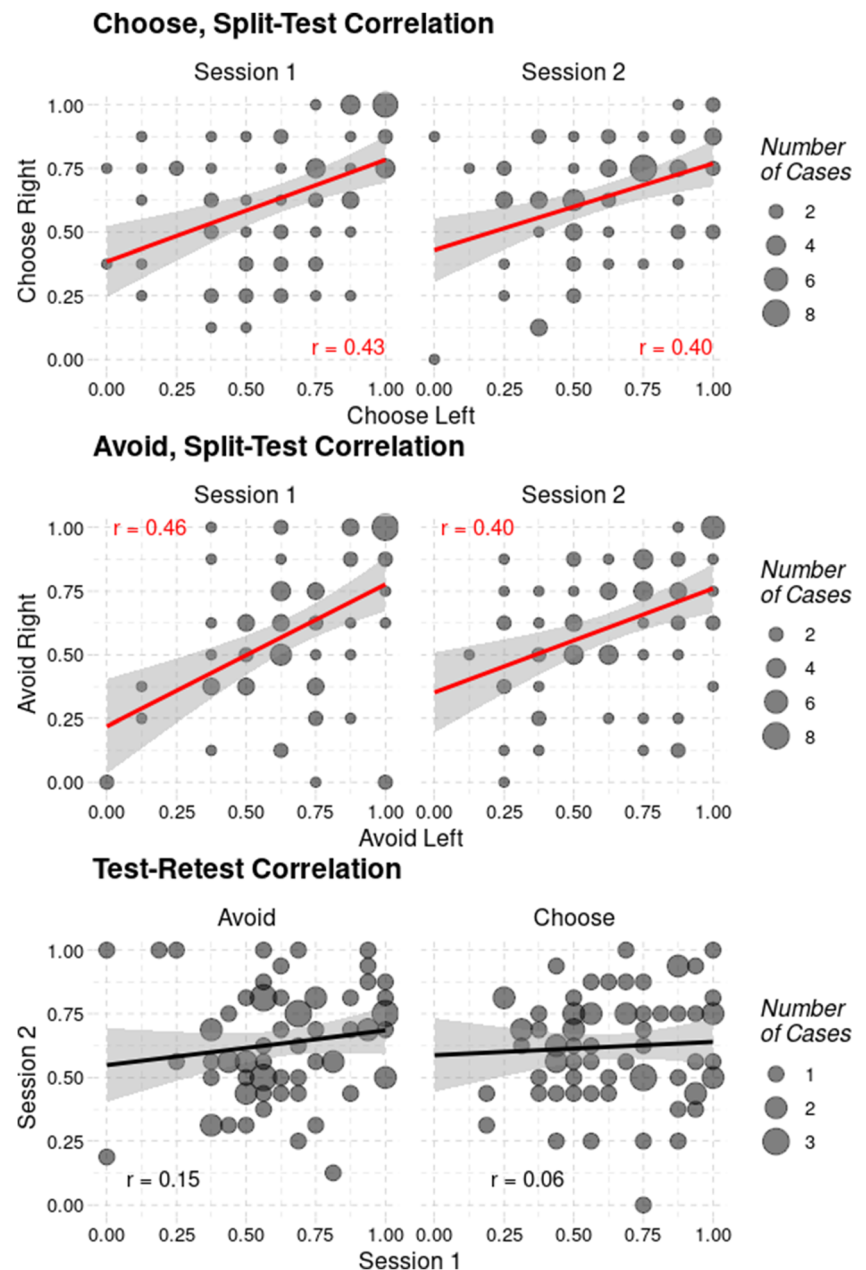
Test-Retest Reliability Then, we examined the test-retest reliability across sessions of the same measures. To do so, we simply computed Choose and Avoid canonically (i.e., across all relevant pairs) for the two sessions. In contrast to the split-test results, no significant correlation was found for either Choose [$r(71) < 0.10, p > 0.60$] or Avoid [$r(71) = 0.15, p > 0.20$] across sessions (Fig. 3).

Intraclass Correlation Coefficient Finally, for each of the measures of interest, we also calculated the intraclass correlation coefficient (ICC: Shrout & Fleiss, 1979). Conceptually, ICC measures the proportion of variance in the measures of interest against the total variance, and is used as an assessment of the consistency of quantitative measurements between sessions. In our study, the variance of interest is between different participant measurements *P* of the same variables (Choose or Avoid) across left/right presentation or sessions, and the total variance is due to the different measurements *M*, the error *E*, and the individual participants *P*. Thus:

$$ICC = \sigma^2_P / (\sigma^2_M + \sigma^2_E + \sigma^2_P)$$

In practice, the variances are calculated by separating the contribution of the Measurement and Participant factors through an ANOVA (Shrout & Fleiss, 1979). Note that, because, in this case, the numerator requires computing a difference between mean sums of squares, the ICC measure can take any negative or positive real values (Shrout & Fleiss, 1979) and, unlike correlation coefficients, is not constrained

Fig. 2 Split-test and test-retest reliability of Choose and Avoid accuracies. Red lines represent significant, and black lines represent non-significant, correlations



to be between 0 and 1. Higher and positive ICC values denote greater reliability.

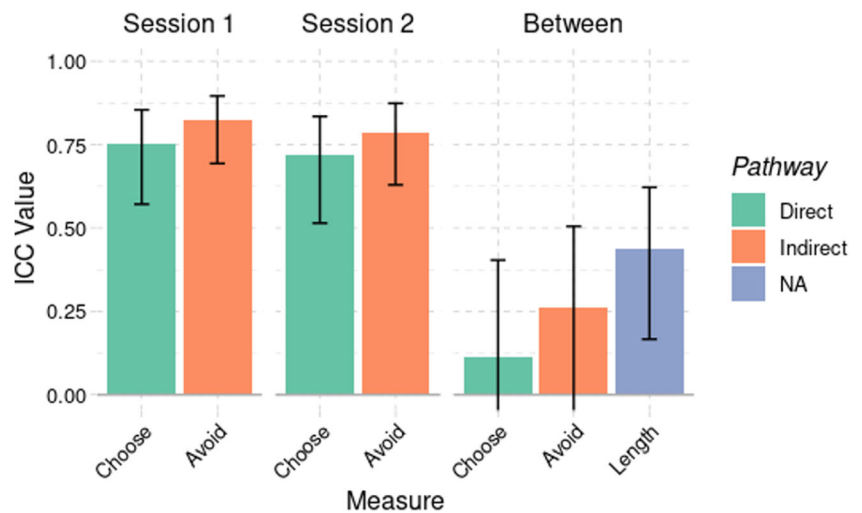
Because the ICC values are computed from an ANOVA, it is important to report the type and model of the analysis. In this case, we applied the ICC(2,k) model of Shrout and Fleiss (1979), which is appropriate for variables that average over different items (the individual choices) and for studies of test-retest reliability, which require absolute agreement (Koo & Li, 2016). As shown in Fig. 3, although the ICC values for Choose and Avoid between sessions were greater than zero, they were also markedly inferior to their split-half counterparts and both values fell below the 0.40 threshold indicated by Cicchetti (1994) as “poor” reliability. We also calculated ICC for

Training Length (that is, numbers of blocks of training) across sessions to serve as a comparison against Choose and Avoid, since a longer training phase likely indicates slower learning and, therefore, should be consistent across sessions. As it shows on the Fig. 3, Choose and Avoid between sessions were markedly less consistent than Length.

Summary

An analysis of the Choose and Avoid measures in the PSS task has yielded somewhat contrasting results. Across sessions, both measures show very poor reliability as indexed by both Pearson correlations and ICC values. This finding is

Fig. 3 Intraclass correlation coefficients of Choose (green) and Avoid (orange) measures within (left and center panel) and between sessions (right panel). Choose and Avoid showed poor reliability across sessions, compared with Length (blue). Error bars represent 95% confidence intervals; colors represent the basal ganglia pathway the variable is supposed to measure



in line with the reports of Grogan et al. (2017) and Baker et al. (2013), which called into question the original results by Frank. On the other hand, the same measures had good reliability scores within each session ($r > 0.4$, $ICC \geq 0.5$), suggesting that the two measures were not intrinsically unreliable.

At least three possible explanations exist for these findings. One is that Choose and Avoid do not index any underlying stable feature of a participant’s biology (such as the relative strengths of their basal ganglia pathways) but some other characteristic that is reliable only within a single session. This could be, for example, the case of transient mental states such as fatigue. A second possibility is that Choose and Avoid do, indeed, index their hypothetical counterparts, but their measure is strongly affected or mediated by other factors that we were not capable of controlling experimentally, such as caffeine or motivation. Finally, another hypothesis is that Choose and Avoid might be intrinsically noisy and poor indicators of the underlying basal ganglia activity. Imagine, for instance, a participant with equally strong direct and indirect pathways; in one session, the direct pathway might be randomly selected early on, be reinforced by positive feedback, and end up dominating throughout, while in a different session the same could happen for the indirect pathway. This would lead to opposite Choose and Avoid scores simply in virtue of the interactions between the underlying biology and the task dynamics.

To distinguish between these three hypotheses, we decided to take a computational approach. First, we examined the performance of an existing, biologically-plausible model of the PSS task. Then, we applied Bayesian methods to estimate the most likely underlying model parameters for each participant, and examined whether such idiographic parameters (whose value is informed, in a Bayesian fashion, by the models known dynamics) would exhibit greater reliability than the raw behavioral measures.

Computational Model and Maximum Likelihood Parameter Estimation

The model used in this study was originally published in Stocco (2018) and its code made available online.¹ Although other formal models exists that incorporate the dynamics of the two basal ganglia pathways (Collins & Frank, 2014; Rice & Stocco, 2017), this model has the advantage of having been developed using the Adaptive Control of Thought - Rational (ACT-R) architecture, which is currently the most common cognitive architecture in use (Kotseruba & Tsotsos, 2018). Because ACT-R is an *integrated* cognitive architecture, this model assumptions and its parameters can be easily interpreted in relation to other cognitive mechanisms and their underlying neural circuits (Anderson, 2007; Anderson et al., 2008).

Similar to other architectures, ACT-R contains dictionary-like structures called “chunks”, which are used to represent static information like semantic memory (“a dog is walking”), perceptual inputs (“red rectangle on the left”), or motor commands (“press the green button”). These chunks are then placed into specialized modules (such as “vision”) where they become accessible to procedural knowledge (represented as “production rules” or “productions”) to carry out cognitive and motor operations. Only one production is selected at any given time and only one production is allowed to fire; this production is selected amongst all of the possibly competing rules on the basis of its relative *utility*, a scalar value that represents the estimated future rewards generated by their applications and is learned through a reinforcement learning algorithm.

The model by Stocco (2018) assumes that PSS task performance relies entirely on procedural knowledge. This assumption is justified by a number of considerations, the first of which concerns the reinforcement learning nature of the task and its reliance, at the biological level, on the basal ganglia. In

ACT-R, the basal ganglia are associated with the procedural module, which manages and controls procedural rules (Anderson, 2005; Anderson et al., 2008; Stocco & Anderson, 2008) and learns the best action selection policy through reinforcement learning. In addition, and as previously noted, the PSS task was explicitly designed to exclude the possibility that participants were relying on declarative knowledge. In fact, experimental results show that task performance is consistently affected by factors that influence basal ganglia function (such as dopamine; Frank et al., 2004; or the expression of striatal dopamine receptors: Frank et al., 2007b) and not affected by manipulation on the formation of declarative memories (Frank et al., 2006). Therefore, Stocco's (2018) ACT-R model implemented a procedural-only approach for the PSS task.

In the model, different procedural actions compete for execution, reflecting the role of the basal ganglia in action selection (Gurney et al., 2001; Houk et al., 2007). Once the stimuli on the screen are encoded as a visual chunk, all the productions that match the current stimuli compete for execution. To capture the competition between the basal ganglia's direct and indirect pathways, each of the productions is divided into a pair of productions, one reflecting the contribution of the direct pathway and one reflecting the indirect pathway. Specifically, the model uses opposing and competing "Choose" and "Avoid" productions for each stimulus (Fig. 4). Therefore, for each production "Choose A", a paired production "Avoid A" is also created to perform an opposite action (*B* in Fig. 4). Note that while a Choose production is created for a certain stimulus, its corresponding Avoid production will select any other option available except that stimulus. For example, "Avoid A" will result in selecting any other stimulus that was paired with *A* on the screen. The two sets of production have complementary effects on the decision process.

Thus, for every pair of stimuli, four possible productions can be performed, corresponding to a pair of "Choose" and "Avoid" actions for each stimulus. The competition between these productions is resolved by comparing each production rule's respective *utility*, a scalar quantity that approximates each production's intrinsic value. At each cycle, the production rule with the highest utility is chosen. To implement softmax action selection, random noise is added to the utility values of the competing production rules. The amount of noise is normally distributed around zero, with a variance σ^2 controlled by a free parameter *s* according to the equation:

$$\sigma^2 = (\pi s^2)/3 \quad (1)$$

A production's utility reflects its historic record in leading to rewards. Specifically, after each reward *R*, the utility U_p^t of

a production *p* at time *t* is adjusted through the following reinforcement learning equations:

$$\begin{aligned} U_p^t &= U_p^{t-1} + \alpha (D_1 R^t - U_p^{t-1}) \text{ for "Choose" productions} \\ U_p^t &= U_p^{t-1} + \alpha (D_2 R^t - U_p^{t-1}) \text{ for "Avoid" productions} \end{aligned} \quad (2)$$

The model parameter α represents the learning rate in reinforcement learning, that is, how much each production's utility *U* is adjusted for after each feedback. The effect of the magnitude of a reward *R* at time *t* is further modulated by D_1 and D_2 , which model the density of dopamine D1 and D2 receptors in the basal ganglia's direct and indirect pathways, respectively. Thus, D_1 and D_2 represent the unobserved quantities that the Choose and Avoid measures are purported to operationalize. The idea of having D_1 and D_2 modulating the *size* of the reward (as opposed to the learning rate, as in O'Reilly & Frank's 2006 model) comes from the notion that, as the number of receptors on each pathway increases, the magnitude of dopamine's effect on the corresponding pathway is magnified. It implies that the size of prediction error (which dopamine conveys: Schultz et al., 1997) would be considered as larger, indicating a greater distance between the previous utility *U* of a production *p* and the actual reward *R* at time *t*. The easiest way to convey this greater distance is to increase the value of R^t .

Thus, the functioning of the model is governed by four parameters only: α , *s*, D_1 , and D_2 (Fig. 4, Eqs. 1 and 2). The original paper (Stocco, 2018) provides values for the learning rate α and noise parameter *s* for the general population, as well as the distribution of values of D_1 and D_2 that capture the observed variability in healthy individuals. In this study, we used the values of $\alpha = 0.018$ and *s* = 0.1 (which were fit to the healthy control data in Stocco, 2018), and parametrically varied the values of D_1 and D_2 from 0 to 2 in increments of 0.05 (the same range was used in Stocco, 2018, to capture individual differences in the PSS task). For each combination of D_1 and D_2 parameter values, the model was then run 250 times, and the probability distributions of each combination of Choose and Avoid measures were recorded. A total of $41 \times 41 \times 250 = 420,250$ simulations were run.

Likelihood Distributions

An inspection of the likelihood distributions of the model's performance provided a first insight into the reasons for the poor test-retest reliability of the PSS task measures. Under certain combinations of parameters, the model tends to converge on the same estimates of Choose and Avoid; this result is represented by a likelihood distribution with a unique global maximum (Fig 5, left). However, under most combinations, the likelihood distribution did not have a single maximum,

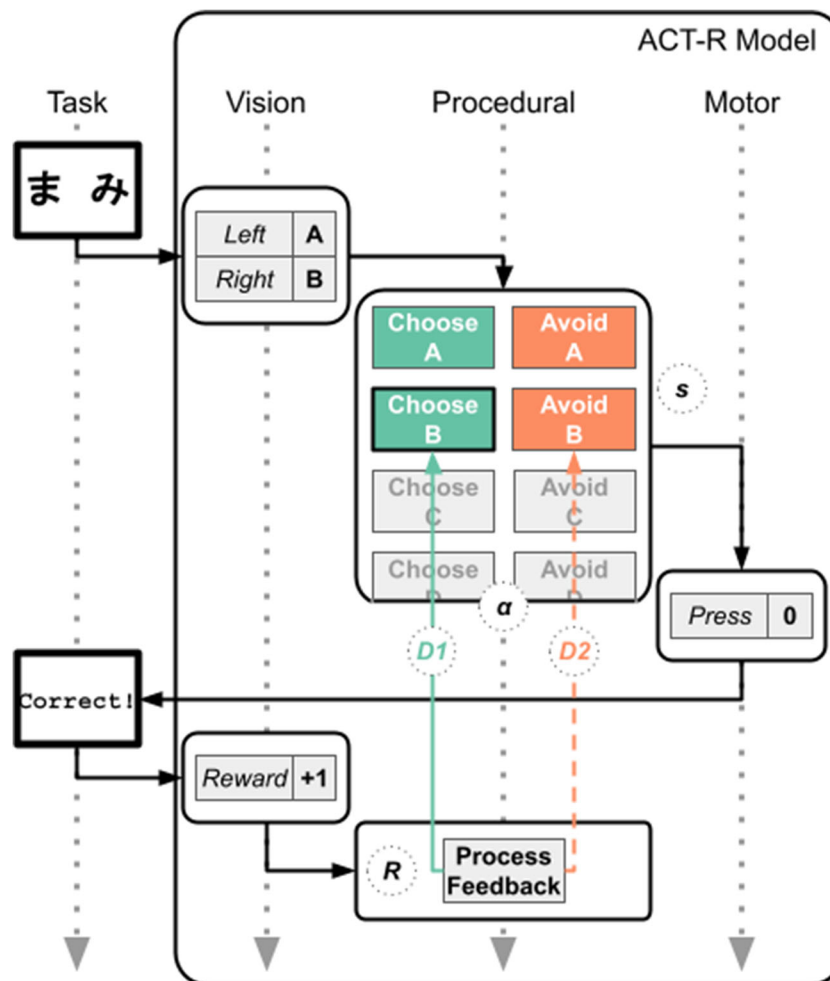


Fig. 4 Overview of the ACT-R model of the PSS task (as described in Stocco, 2018) performing the PSS task. The figure depicts the passage of time (dotted grey arrows) in a sample trial of the task. Rounded boxes represent the activity of three modules (visual, producedural, and motor) over time; squares represent example data structures used by each module (chunks for the visual and motor modules, production rules for the procedural module); circles represent the points at which the model parameters (α , D_1 , D_2 , s , and R : Eq. 1 and 2) affect computation. In this example, the model is presented with a *A-B* pair; the corresponding productions

for stimuli *A* and *B* match the visual stimuli (visually depicted as being colored vs. grey) and compete for selection; the winning production “Choose *B*” is selected (shown as a thick contour), triggering the motor response “Press 0” in the Motor module. The task’s visual feedback is then translated into a reward value and processed by the procedural module, causing an update in the utility of “Choose *B*” through the modulation of the reward quantity R by D_1 and α (Eq. 2). If one of the “Avoid” productions were chosen, the effect of reward would have been modulated by D_2 (orange dotted line) instead of D_1

and the model would move towards different values for Choose and Avoid in different runs. (Note that, since there is a finite number of pairs in the testing phase, the values of Choose and Avoid are discrete, and the multiple peaks in Fig. 5 do not represent an approximation due to the discretization of continuous variables). This suggests that the poor reliability of Choose and Avoid accuracies might be due to the nature of the task and the ways the measures are calculated.

Maximum a Posteriori (MAP) Parameter Estimation

The simulations described in the previous section provide an estimate of the likelihood of observing a particular behavioral

outcome Y (that is, a combination of Choose and Avoid values) given θ (that is, a combination of values for the D_1 and D_2 parameters). In addition to these likelihood estimates, we were interested to explore whether the model’s simulations could be used to estimate reliable values for D_1 and D_2 (the unobservable parameters that govern learning rates in the two pathways) from the observable behavioral measures (Choose and Avoid). To do so, we fitted the model to each individual participant using a Bayesian Maximum A Posteriori (MAP) procedure. In Bayesian statistics, a MAP is the estimate of the maximum likelihood of an unobservable quantity on the basis of both the empirically observed data and a prior hypothesis about the distribution of that quantity. In our case, the procedure was used to recover the most likely values of θ (D_1 and

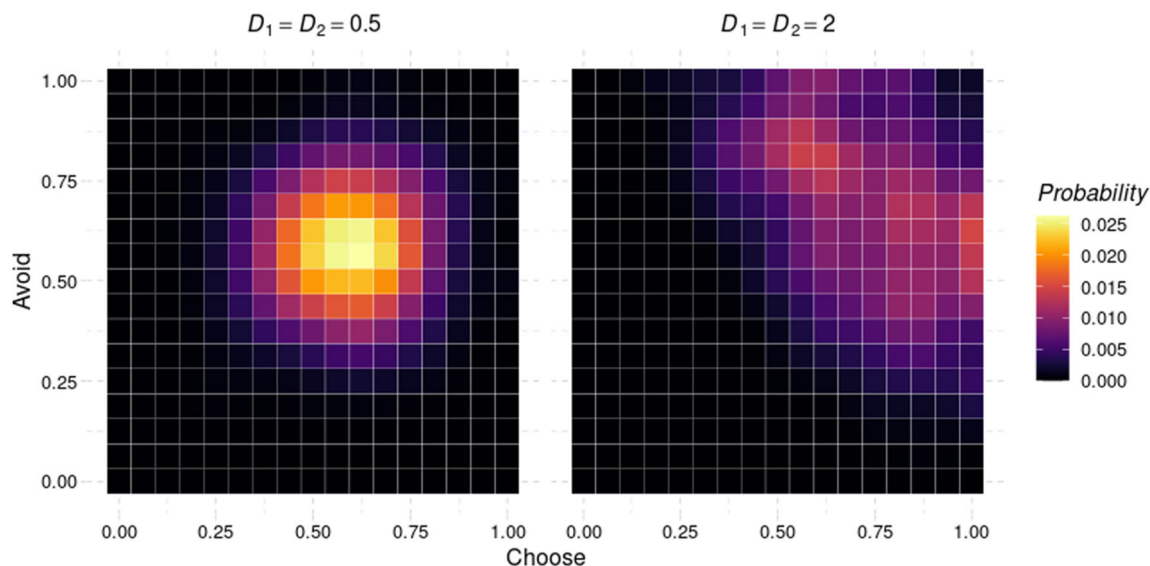


Fig. 5 Variability in the likelihood of possible results as a function of different parameter values θ . When $D_1 = D_2 = 0.5$ (Left), the model converges on a single global maximum (Choose = 0.6125, Avoid =

0.6125). However, when $D_1 = D_2 = 2.0$ (Right), multiple possible results (i.e., maxima) are equally likely. Colors represent probability densities for each Choose/Avoid accuracy combination

D_2) given the observed Choose and Avoid values of a given participant (Y), that is, $\operatorname{argmax} P(\theta | Y)$. Using Bayes theorem, this quantity can be rewritten as:

$$P(\theta | Y) = \operatorname{argmax} [P(Y | \theta) \times P(\theta) / P(Y)].$$

The likelihood value $P(Y | \theta)$, that is, the distributions of Choose and Avoid accuracies given pairs of D_1 and D_2 parameters, can be directly computed from the model simulations (Fig. 5). To estimate the parameter priors $P(\theta)$, we followed the following logic. First, we modeled the probability distribution of each parameter value as a normal distribution $N(\mu, \sigma)$ with mean $\mu = 1$ and $\sigma = 0.5$. This captures the finding that the values for D_1 and D_2 that best represent variability among healthy participants vary between 0.5 and 1.5, with 1 being the population mean (Stocco, 2018). As a final step to define the priors, we need to define joint probability distribution of D_1 and D_2 . In turn, this requires providing an estimate of the degree of the correlation of these two parameters in the general population. The existence of a correlation between the two parameters is suggested by three layers of evidence. First, in the basal ganglia, studies (Sian et al., 1999) showing synergic biochemical and behavioral effects associated with combined administration of D1 and D2 agonists suggest that the distribution of D1 and D2 receptors is not independent. Second, the activity in both receptors is driven by a common force (the release of dopamine). Finally, the recorded activity of the two pathways is anticorrelated, suggesting that a greater D_1 means more excitation on the indirect pathways, and greater D_2 means more inhibition on the direct pathway. Although this evidence does point to the existence of a correlation, it does not precisely define its magnitude.

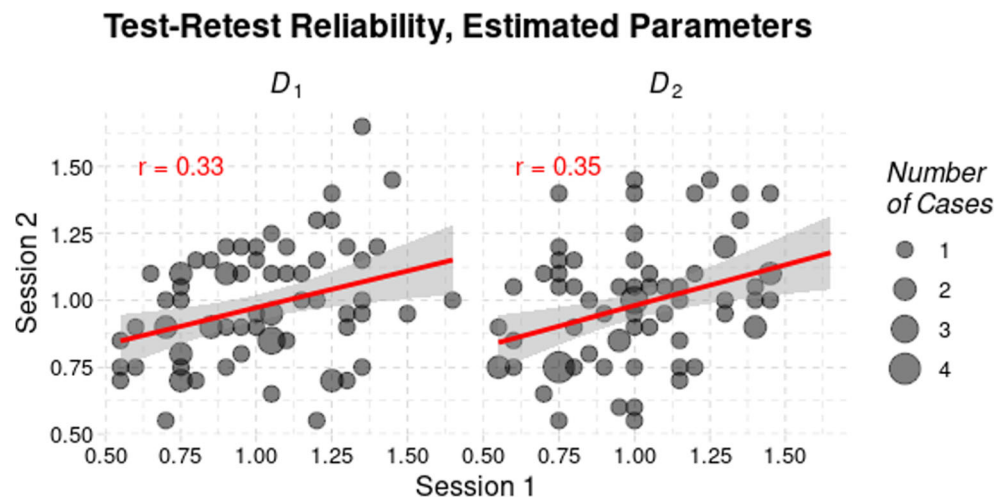
Here, we decided to set the correlation between the two distributions to the agnostic value of $r = 0.5$, which is exactly in the middle of the range between 0 (complete independence) and 1 (complete dependence). This estimate is undoubtedly imprecise, and could be improved upon in the future.

Taken together, the use of likelihoods and joints priors might overcome the problem of parameter identifiability (Fig. 5). While it remains true that, given the same initial conditions and values for D_1 and D_2 , the model can still produce different outcomes, the MAP procedure greatly reduces the possible range of plausible initial values. Thus, we expect that the MAP estimates of D_1 and D_2 for each participant would have a higher reliability than the simple Choose and Avoid accuracies. With both the likelihood and the joint probability distribution in place, we proceeded to test this prediction.

Test-Retest Reliability

After calculating the MAP parameter estimates for each participant, we applied the same test-retest reliability analyses that were used for the behavioral measures to the individual parameter values. In contrast to our behavioral findings (Fig. 2), we found statistically significant Pearson correlations across sessions for both D_1 [$r(71) = 0.33$, $p < 0.005$] and D_2 [$r(71) = 0.35$, $p < 0.003$; Fig. 6]. Furthermore, the correlation values for the model parameters were either significantly higher or marginally significantly higher than the correlation values for the corresponding behavioral measure, as shown by the difference of their correlation coefficients, converted into Z values using Fisher's transformation (D_1 vs. Choose: $Z = -1.71$, $p = .04$; D_2 vs. Avoid: $Z = -1.31$,

Fig. 6 Correlation between MAP estimates of the underlying D_1 and D_2 model parameters across sessions for all participants. Red lines indicate significant correlations



$p = .09$, one-sided). Similar results were found for the corresponding ICC values, with the values for D_1 (ICC = 0.49) and D_2 (ICC = 0.51) being more than twice as large as the corresponding values for Choose and Avoid, and within Cicchetti’s (1994) range of “fair” reliability (Fig. 7). As a comparison, Fig. 7 also includes Training Length, with its value being 0.44. Thus, the ICC values for D_1 and D_2 are even larger than the corresponding value for Length, implying a remarkably improved consistency between sessions.

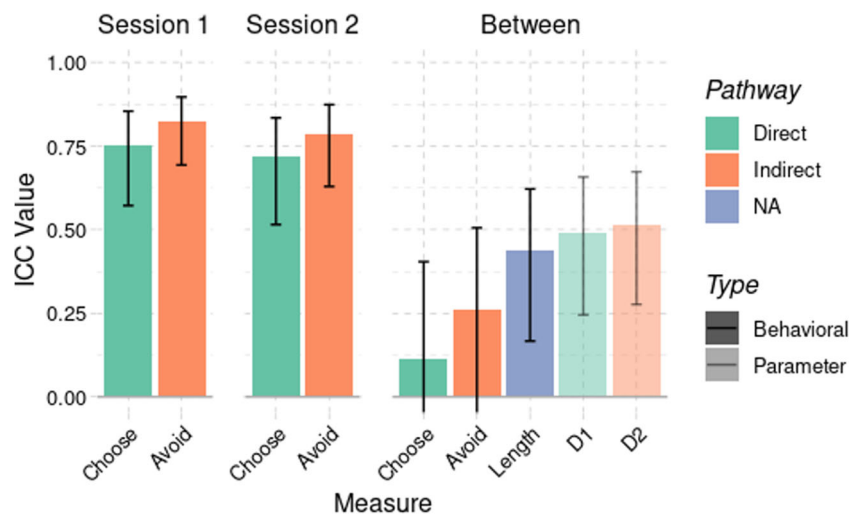
Discussion

An analysis of the behavior of a biologically plausible computational model of the PSS task (Stocco, 2018) suggests that the lack of test-retest reliability for the Choose and Avoid measures is due to the one-to-many relationship between the underlying neural parameters and the observed behavior. Specifically, certain combinations of D_1 and D_2

parameters result in equal probability densities for different values of Choose or Avoid, making the correlations between different runs intrinsically noisy.

However, the intrinsic noise in the Choose and Avoid measures does not rule out the possibility of recovering important information from behavior. The key to this insight is that, while multiple outcomes might be equally likely, given an initial set of biological parameters, not *all* possible behavioral outcomes are. Using an exhaustive set of simulations from the model, we were able to apply a Maximum A Posteriori parameter estimation procedure to map each behavioral outcome to the most likely set of model parameters that could lead to the observed combination of Choose and Avoid accuracies. These idiographic parameters are more clearly interpretable and represent a more direct operationalization of their hypothetical biological substrate; perhaps because of this they proved to be superior to the pure behavioral measures in terms of reliability.

Fig. 7 A comparison of the intraclass correlation coefficient values of the behavioral measures (solid) and the MAP estimates of the underlying model parameters (transparent) that index the function of the two pathways, with Length (blue) as a comparison. Error bars represent 95% confidence intervals; colors represent the basal ganglia pathway the variable is supposed to measure



Experiment 2

The results of our computational model analysis suggest that it is possible to overcome some of the limitations of the behavioral measures of the PSS tasks. To further ensure our model-guided way is more reliable in recovering individual differences, a follow-up experiment was done using a modified version of the PSS task. In this version, the original Hiragana characters were replaced with a new set of characters from an unfamiliar, extinct language—Suyat. We modified the stimuli also under the concern that, as the original PSS task can be only presented to non-Japanese speakers, it might jeopardize the generalizability of the task. Unlike the original Hiragana stimuli, these particular Suyat characters were extremely unlikely to have been previously seen by any participants. They also consisted of different arrangements of identical perceptual features (cuneiform strokes), making them more perceptually similar to one another, therefore limiting the preference issue on perceptual discriminability raised by Schutte et al. (2017).

Methods and Materials

Participants Forty-two undergraduate students from the University of Washington (age 18–22, 21 females) were recruited following the same process as Experiment 1, taking part in this experiment in exchange for credit course. Although it was originally planned that the same number of participants as Experiment 1, data collection had to be suspended early due to the end of the academic quarter, and could not be resumed due to the COVID-19 pandemic. All participants completed two sessions of the modified PSS task exactly 1 week apart. As in Experiment 1, the second session always occurred on the same day of week and at the same time of day as the first session. All of the study procedures were approved by the University of Washington Institutional Review Board (Application #00000820). Data from six participants (3 females) was lost to attrition (that is, they did not return to complete the second session), leaving a total of 39 participants in the sample.

Task The modified PSS task used the same code as Experiment 1, but, instead of a Hiragana typeface, it presented the stimuli in a typeface designed to imitate the Suyat script. Suyat, now extinct, is the modern collective name of the indigenous scripts of various ethnolinguistic groups in the Philippines.

All of the remaining task components, such as the duration of each stimulus, the chances of being correct or incorrect for each stimulus, the learning criterion, and the way Choose and Avoid accuracies are measured, were all kept the same. Participants were asked to place their left index finger on

button “1” and right index finger on button “0” of a standard computer keyboard placed in front of them. Pairs of Suyat characters were then shown on the screen with a fixation cross in between each trial. They then pressed the button corresponding with the characters they intuitively think would be correct. Feedbacks were shown on the screen after each selection during the training phase as “Correct!” in blue color or “Incorrect!” in red color. If the participant did not press any button within 6 seconds, “no response detected” in red color was shown on the screen. Again, this was to ensure that the subject was engaging in the task and also to discourage the subject from using explicit methods to remember the patterns rather than learning them through trial and error. Still, a maximum of six repetitions of the training phase was followed by a testing phase where Choose and Avoid accuracies were measured.

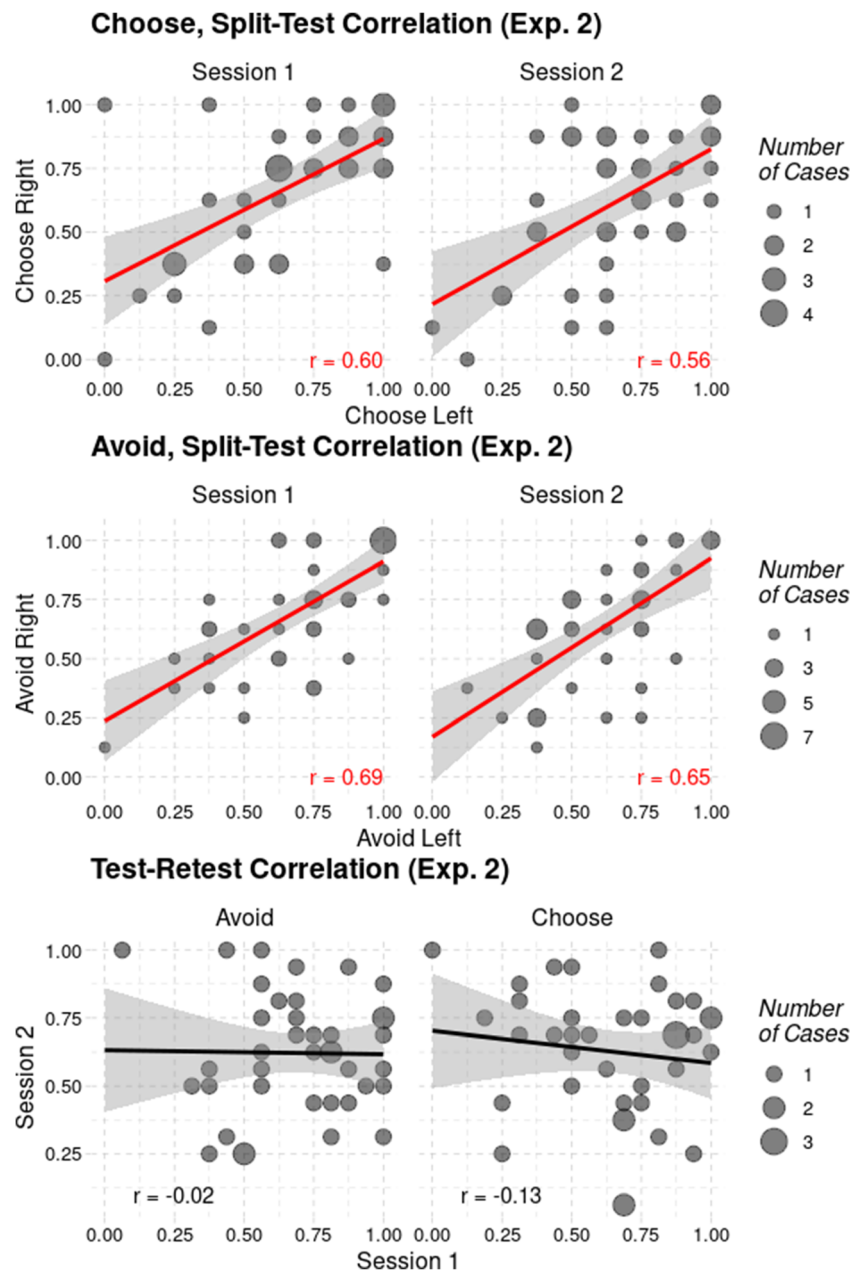
Results

Split-Test Reliability First, we examined the split-test reliability of these new measures in the same way we did in Experiment 1. We calculated separately the values of the two main variables, Choose and Avoid, for different pairs of stimuli, depending on whether *A* and *B*, were presented on the left (e.g., “*A-C*”, “*B-C*”, etc.) or on the right (e.g., “*C-A*”, “*C-B*”, etc.). Again, these measures were called Choose Left, Avoid Left, Choose Right, and Avoid Right, respectively. The Pearson correlation between the Left and Right version of each measure was calculated. As shown in Fig. 8, the split-test correlation coefficients of Choose and Avoid were significant in both sessions. Specifically, we found a positive correlation of Choose Left and Choose Right in Session 1 [$r(39) = .60, p < .001$] and Session 2 [$r(39) = .60, p < .001$], and between Avoid Left and Avoid Right in Session 1 [$r(39) = .70, p < .001$] and in Session 2 [$r(39) = .65, p < .001$].

Test-Retest Reliability Then, we examined the test-retest reliability across sessions of the same measures. In contrast to the split-test correlations, and in line with the results of Experiment 1 no significant correlation was found for either Choose [$r(39) = -.02, p > .44$] or Avoid [$r(39) = -.13, 0, p > .92$] across sessions (Fig. 8).

Intraclass Correlation Coefficient Finally, for each of the measures of interest, we also calculated the ICC scores. This time, the ICCs for the modified version are remarkably worse than those of the original version. As it shows in Fig. 9, the ICC values for Choose and Avoid between sessions are negative. As noted above, while negative ICCs are possible, they are interpreted as implying that true intraclass correlation is effectively zero, meaning that two members chosen randomly from any class vary almost as much as any two randomly chosen members of the whole population (Taylor, 2010). As in

Fig. 8 Split-test and test-retest reliability of Choose and Avoid accuracies. Red lines represent significant, and black lines represent non-significant, correlations



Experiment 1, the Length of the training session was also included to serve as a comparison against Choose and Avoid.

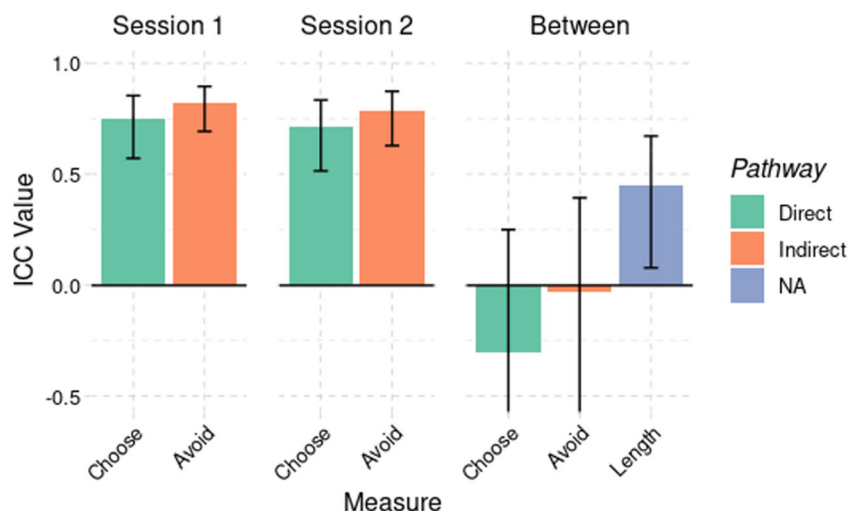
Maximum A Posteriori (MAP) Parameter Estimation

An analysis of the Choose and Avoid measures in the modified PSS task yielded the same pattern of results of Experiment 1. While both measures showed reasonable split-test reliability, they exhibited very poor test-retest reliability as indexed by both Pearson correlations and ICC values. Thus, the results show poor reliability for Choose and Avoid regardless of the perceptual saliency or

discriminability of the stimuli associated with more and less rewarding outcomes, respectively. To examine the performance of cognitive modeling and the reliability of idiographic parameters in the context of the new visual stimuli, we applied the same Bayesian methods of Experiment 1 to estimate the most likely underlying model parameters for each participant.

After estimating the idiographic parameters of the modified PSS task for each participant, we applied the same test-retest reliability analyses that were used for both Experiment 1 and the behavioral measures to the individual parameter values. In contrast to our current behavioral findings (Fig. 8), and in line with our previous parameter findings (Fig. 6), we found

Fig. 9 Intraclass correlation coefficients (ICC) results. Reliability of Choose (green) and Avoid (orange) measures within (left and center panel) and between sessions (right panel). Choose and Avoid showed poor consistency across sessions, compared with Length (blue). Error bars represent 95% confidence intervals; colors represent the basal ganglia pathway the variable is supposed to measure



Pearson positive correlation across sessions for both D_1 [$r(39) = 0.29, p = .073$] and D_2 [$r(39) = 0.29, p = .072$; Fig. 9]. Because of the reduced sample size of Experiment 2, our Pearson correlation coefficients were only marginally significant; however, they are entirely comparable, in magnitude, to those found in Experiment 1, and, like in Experiment 1, they remain superior to the corresponding behavioral measures (Choose vs. D_1 : $Z = -1.4, p = .08$; Avoid vs. D_2 : $Z = -1.83, p = .03$). Similar results were found for the corresponding ICC values, with the values for D_1 (ICC = 0.43) and D_2 (ICC = 0.42) being exceedingly larger than the corresponding values for Choose and Avoid (Fig. 10). We also included Length as a comparison, with its value being 0.52. Although for this modified PSS task, ICC values for D_1 and D_2 did not exceed that of Length, they still fall within Cicchetti's (1994) range of "Fair" reliability and indicate an improved consistency between sessions (Fig. 11).

Discussion

In this paper, we have presented evidence that computational cognitive models can be used to recover interpretable and reliable parameters from noisy behavioral data, making up for discrepant findings. More specifically, the study examined the reliability of two commonly used measures in the PSS task and suggested that recovering the underlying parameters by fitting cognitive models could be used as an alternative to simple metrics derived from behavioral choice data.

The study consisted of two experiments, each of which with two parts. During the first experiment, we tested the reliability of the original PSS task's most important measures, Choose and Avoid accuracies, and showed that their test-retest reliability is poor. Across two sessions 1 week apart, both Choose and Avoid values were uncorrelated within participants. An existing ACT-R model, capable of simulating the competitive dynamics of the two basal ganglia pathways that

Fig. 10 Correlation between MAP estimates of the underlying D_1 and D_2 model parameters across sessions for all participants. Orange lines indicate marginally significant correlations

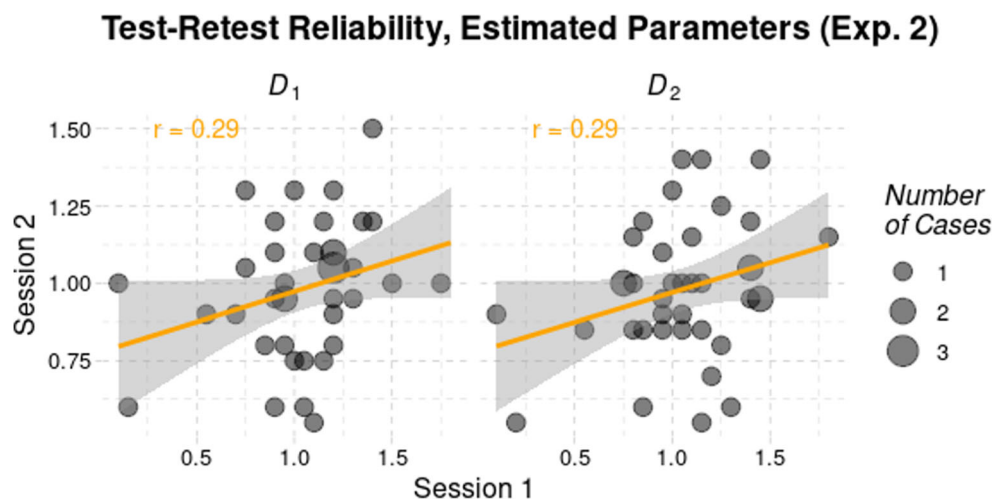
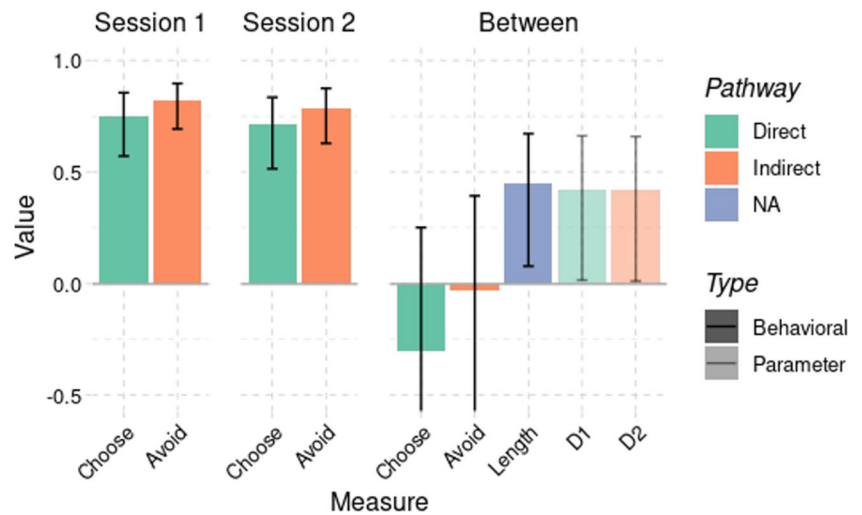


Fig. 11 A comparison of the intraclass correlation coefficient (ICC) values of the behavioral measures (solid) and the MAP estimates of the underlying model parameters (transparent) that index the function of the two pathways, with Length (blue) as a comparison. Error bars represent 95% confidence intervals; colors represent the basal ganglia pathway the variable is supposed to measure



jointly drive the reinforcement learning process, was then used to generate probability distributions of Choose and Avoid value for a broad range of underlying model parameters. Using a Maximum A Posteriori procedure, the most likely values for two of critical model parameters, D_1 and D_2 , were then obtained for each of our participants. As a result, both D_1 and D_2 show significant Pearson correlations and greater intraclass correlation coefficients across sessions than the original Choose and Avoid measures. Next, to address a limitation on generalizability, a modified version of the PSS task using Suyat instead of Hiragana characters was introduced. The same model-fitting procedures and reliability tests for behavioral measure and parameter measure were conducted, and the results replicated the findings of the first experiment, with D_1 and D_2 parameters showing greater reliability across sessions than Choose and Avoid measures.

It should be noted that, though the reliability and the replicability of the behavior results of the PSS task has been under debate, it is no doubt the task itself has been widely used on indexing cognitive processes. Despite studies (mentioned above) showing poor replicability or poor reliability across sessions, people also found the PSS greatly useful and dependable on indexing the correlation between DA level and the sensitivity to reward learning and punishment avoidance. McCoy et al. (2019) suggests that dopamine medication reduced negative (but not positive) outcome learning rates. Similar findings have been reported by Maril et al. (2013) that there are opposite medication effects on positive and negative outcome learning rates, but only in patients with left hemisphere dopamine depletion. This idea is also supported by Kobza et al. (2012) and Weismueller et al. (2018), both of which showed a strong negative learning bias for PD patients while off medication. Thus, on a more operational level, our results imply that one needs to exert caution in

interpreting data from individual subjects in the PSS task, and recovering idiographic parameters is a new analysis tool, here shown to be reliable, when conducting individual subject analysis with behavioral results.

Limitations

While our results are encouraging, a number of limitations need to be acknowledged. First, our two experiments did not have the same sample size. Our data collection had to end earlier than planned due to the COVID-19 pandemic, resulting in a much smaller sample for the second experiment. As a result, while the size of the Pearson and intraclass correlation coefficients were similar across both experiments, some of the correlations in the second experiment did not reach the canonical statistical significance threshold of $p < .05$. However, this limitation could be easily overcome in future, if and when data collection can resume safely, or in further replication experiments.

A second limitation is that our analysis was restricted to young, healthy adults. In contrast, a sizable amount of PSS studies focused on patient populations (e.g., Cicero et al., 2014; Frank et al., 2007a; Waltz et al., 2007). Thus, a stronger demonstration of the reliability of this approach would be if (a) reliable parameters could be recovered from patients populations, and (b) the value of the recovered parameter were consistent with what would be expected from the etiology of the disorder. In fact, since the original model (Stocco, 2018) had already been fit the group data from PD patients) originally published in Frank et al., 2004), a unique opportunity exists to test both the reliability as well as their expected values of D_1 and D_2 parameters in this population.

A third limitation concerns the number of model parameters that were manipulated for the estimation procedure. Here, we have focused exclusively on D_1 and D_2 , while keeping α

and s at the values recommended in Stocco (2018). However, a complete analysis of individual-level performance would also require fitting α and s to each participant. Being this a preliminary study, we considered that expanding the number of parameters would pose a significant burden the time needed to run simulations and calculate the likelihood distributions. To further complicate the matter, the parameters are interdependent. For example, both learning rate and D_1/D_2 jointly determine the magnitude of the utility update (Eq. 2). Additionally, because the only behavioral outcome is a decision, it is difficult to separate the effects of noise (which alters the decision process) from the effects of learning (which precedes decision, and must be inferred). On the other hand, we do recognize that increasing the number of parameters will expand the range of behaviors that could be modeled, and their inter-dependency could be mitigated by the possibility of using additional behavioral measures to inform the MAP procedure. For example, the length of the training phase would likely be related to individual differences in the learning rate a , and could be used to constrain it.

A fourth limitation concerns the parameter recovery process. Here, we implemented a Bayesian Maximum A Posteriori procedure to identify the most likely parameters, given the participant's data and the thousands of simulations we had run with the model. However, there are other possible ways to proceed. For example, parameters could be identified using a gradient descent algorithm (such as the Nelder-Mead method) instead of a Bayesian approach. It should be noted, however, that gradient descent methods are sensitive to the presence of local or multiple minima, which do occur in the distribution of Choose and Avoid values (see Fig. 5), while our method, while computationally intensive, has the advantage of exploring the full range of parameter values.

A fifth limitation concerns the degree to test-retest reliability that is afforded by the model-based parameter-recovery procedure. While a considerable improvement over raw behavioral data, it might be argued that the values of Pearson and intra-class correlation scores were not as high as one could hope. Ultimately, because the same initial conditions could lead to different outcomes (Fig. 5, right panel), the function mapping between parameter values and Choose and Avoid accuracies is not invertible, and this provides an upper limit to the degree to which “true” parameter values can be measured. Still, we believe that the results reported here can be further improved upon. As noted above, it is possible that the precision of our estimate could be increased by fitting the learning rate in addition to D_1 and D_2 . The estimates could also be made more precise by taking into account trial-by-trial choices, thus increasing the resolution of the observations the model is fit to (Daw, 2011).

Finally, the Suyat stimuli in our modified PSS task in the second experiment could suffer from the same concerns that Schutte et al. (2017) pointed out for the Hiragana characters—

namely, that performance in the PSS task is affected by participant's perceptual preferences for certain stimuli. Because of their reduced familiarity, we suspect that this problem is reduced in Suyat characters. Furthermore, this limitation does not invalidate our finding that cognitive modeling can be used to recover reliable individual differences, but solely points to a future direction when increasing the generalizability of the PSS task. As shown in our inspection of the likelihood distributions of the model's performance (Fig. 5 right), the poor reliability is largely due to multiple equally likely results under certain sets of parameters.

Implications for Future Studies

These limitations notwithstanding, with the PSS task as a starting point, our results do have significant implications for future research. First and foremost, these results suggest that using cognitive models might be a more reliable way of dealing with behavioral data, as idiographic parameters better capture stable individual traits that possess higher test-retest reliability. Idiographic model parameters had been previously proposed to generalize or predict behaviors across tasks; for instance, Lovett and colleagues were able to estimate attentional spreading activation from a working memory task, and use it to predict performance in a second task (Daily et al., 2001; Lovett et al., 2000). This study, however, goes one step further, estimating model parameters that are directly related to biological properties, such as procedural learning rate and the density of dopamine receptors (Stocco, 2018; Stocco et al., 2017). These underlying biological properties have important consequences not only for the Probabilistic Stimulus task, but in many other domains, including response inhibition (Stocco et al., 2017), fluid reasoning (Stocco et al., *in press*) and working memory (Zhang et al., 2007). Therefore, we see a great potential in the application of such model-based approach in the field of computational psychiatry, where it could possibly be used to better quantify and diagnose, for example, the complex symptomatology and downstream effects of basal ganglia dysfunction in Parkinson's Disease. This could be accomplished by estimating idiographic parameters and comparing them against their priors for healthy participants.

Because the model that was used in this study was developed within a general-purpose cognitive architecture, the approach outlined here could be used to incrementally build more comprehensive and detailed computational models of a single person. This is because, in ACT-R, the learning parameters associated with the PSS task (α , s , D_1 , and D_2) maintain the same meaning and interpretation even if they are used in the context of a different task (Stocco et al., 2017). Knowing their values within a simple task (such as the PSS paradigm), therefore, permits us to better estimate parameters that can only be understood using more complex paradigms, such as those measuring working memory capacity.

Of course, for these implications to be realistic, the procedures we have outlined here would need to be expanded upon and tested on other task paradigms, possibly using additional well-established models and recovering other parameters. For this reason, we look forward to seeing future studies in this direction.

Acknowledgements The authors would like to thank Justin Abernethy for help coordinating this study and Megana Boddam for help in collecting the data.

Code Availability The ACT-R code for the model of Fig. 6 as well as the additional Lisp code to run the simulations, all the data from the simulations, and the R code to analyze the data, are available at Cognition and Cortical Dynamics Laboratory's online repository: https://github.com/UWCCDL/PSS_model. The behavioral data collected in Experiment 1 and 2 as well as the R code to analyze the data and to implement the Maximum A Posteriori model fitting procedure can be found at https://github.com/UWCCDL/PSS_Reliability.

References

- Albin, R. L., Young, A. B., & Penney, J. B. (1989). The functional anatomy of basal ganglia disorders.
- Anderson, J. R. (2005). Human symbol manipulation within an integrated cognitive architecture. *Cognitive Science*, 29(3), 313–341
- Anderson, J. R. (2007). How can the human mind occur in the physical universe?. Oxford University Press.
- Anderson, J. R., Fincham, J. M., Qin, Y., & Stocco, A. (2008). A central circuit of the mind. *Trends in Cognitive Sciences*, 12(4), 136–143.
- Baker, T. E., Stockwell, T., & Holroyd, C. B. (2013). Constraints on decision making: implications from genetics, personality, and addiction. *Cognitive, Affective, & Behavioral Neuroscience*, 13(3), 417–436
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological assessment*, 6(4), 284–290
- Cicero, D. C., Martin, E. A., Becker, T. M., & Kerns, J. G. (2014). Reinforcement learning deficits in people with schizophrenia persist after extended trials. *Psychiatry research*, 220(3), 760–764
- Collins, A. G. (2018). The tortoise and the hare: Interactions between reinforcement learning and working memory. *Journal of cognitive neuroscience*, 30(10), 1422–1432
- Collins, A. G., & Frank, M. J. (2014). Opponent actor learning (OpAL): Modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. *Psychological review*, 121(3), 337–366
- Cools, R., Altamirano, L., & D'Esposito, M. (2006). Reversal learning in Parkinson's disease depends on medication status and outcome valence. *Neuropsychologia*, 44(10), 1663–1673
- Daily, L. Z., Lovett, M. C., & Reder, L. M. (2001). Modeling individual differences in working memory performance: A source activation account. *Cognitive Science*, 25(3), 315–353
- Daw, N. D. (2011). Trial-by-trial data analysis using computational models. *Decision making, affect, and learning: Attention and performance XXIII*, 23(1)
- DeLong, M. R. (1990). Primate models of movement disorders of basal ganglia origin. *Trends in neurosciences*, 13(7), 281–285
- Djamshidian, A., Jha, A., O'Sullivan, S. S., Silveira-Moriyama, L., Jacobson, C., Brown, P., Lees, A., & Auerbach, B. B. (2010). Risk and learning in impulsive and nonimpulsive patients with Parkinson's disease. *Movement Disorders*, 25(13), 2203–2210
- Endrass, T., Kloft, L., Kaufmann, C., & Kathmann, N. (2011). Approach and avoidance learning in obsessive-compulsive disorder. *Depression and Anxiety*, 28(2), 166–172
- Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in parkinsonism. *Science*, 306(5703), 1940–1943
- Frank, M. J., O'Reilly, R. C., & Curran, T. (2006). When memory fails, intuition reigns: Midazolam enhances implicit inference in humans. *Psychological Science*, 17(8), 700–707
- Frank, M. J., Samanta, J., Moustafa, A. A., & Sherman, S. J. (2007a). Hold your horses: Impulsivity, deep brain stimulation, and medication in parkinsonism. *Science*, 318(5854), 1309–1312
- Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchison, K. E. (2007b). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proceedings of the National Academy of Sciences of the United States of America*, 104(41), 16311–16316
- Gerfen, C. R., Engber, T. M., Mahan, L. C., Susel, Z. V. I., Chase, T. N., Monsma, F. J., & Sibley, D. R. (1990). D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. *Science*, 250(4986), 1429–1432
- Grogan, J. P., Tsivos, D., Smith, L., Knight, B. E., Bogacz, R., Whone, A., & Coulthard, E. J. (2017). Effects of dopamine on reinforcement learning and consolidation in Parkinson's disease. *Elife*, 6, e26801
- Gurney, K., Prescott, T. J., & Redgrave, P. (2001). A computational model of action selection in the basal ganglia. I. A new functional anatomy. *Biological cybernetics*, 84(6), 401–410
- Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behav Res*, 50, 1166–1186
- Houk, J. C., Bastianen, C., Fansler, D., Fishbach, A., Fraser, D., Reber, P. J., Roy, S. A., & Simo, L. S. (2007). Action selection and refinement in subcortical loops through basal ganglia and cerebellum. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1485), 1573–1583
- Kobza, S., Ferrea, S., Schnitzler, A., Pollok, B., Südmeyer, M., & Bellebaum, C. (2012). Dissociation between active and observational learning from positive and negative feedback in Parkinsonism. *PLoS One*, 7(11), e50250
- Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science*, 273(5280), 1399–1402.
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine*, 15(2), 155–163
- Kotseruba, I., & Tsotsos, J. K. (2018). 40 years of cognitive architectures: Core cognitive abilities and practical applications. *Artificial Intelligence Review*, 1–78
- Lovett, M. C., Daily, L. Z., & Reder, L. M. (2000). A source activation theory of working memory: Cross-task prediction of performance in ACT-R. *Cognitive Systems Research*, 1(2), 99–118
- Maril, S., Hassin-Baer, S., Cohen, O. S., & Tomer, R. (2013). Effects of asymmetric dopamine depletion on sensitivity to rewarding and aversive stimuli in Parkinson's disease. *Neuropsychologia*, 51(5), 818–824
- Mathar, D., Neumann, J., Villringer, A., & Horstmann, A. (2017). Failing to learn from negative prediction errors: Obesity is associated with alterations in a fundamental neural learning mechanism. *Cortex*, 95, 222–237
- McCoy, B., Jahfari, S., Engels, G., Knapen, T., & Theeuwes, J. (2019). Dopaminergic medication reduces striatal sensitivity to negative outcomes in Parkinson's disease. *Brain*, 142(11), 3605–3620
- O'Reilly, R. C., & Frank, M. J. (2006). Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation*, 18(2), 283–328.
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716

- Palminteri, S., Lebreton, M., Worbe, Y., Grabli, D., Hartmann, A., & Pessiglione, M. (2009). Pharmacological modulation of subliminal learning in Parkinson's and Tourette's syndromes. *Proceedings of the National Academy of Sciences*, 106(45), 19179–19184
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, 442(7106), 1042–1045
- Piray, P., Zeighami, Y., Bahrami, F., Eissa, A. M., Hewedi, D. H., & Moustafa, A. A. (2014). Impulse control disorders in Parkinson's disease are associated with dysfunction in stimulus valuation but not action valuation. *Journal of Neuroscience*, 34(23), 7814–7824
- Rice, P. J., & Stocco, A. (2017). Basal ganglia-inspired functional constraints improve the robustness of Q -value estimates in model-free reinforcement learning. In *Proceedings of the 15th International Conference on Cognitive Modeling* (pp. 67–73)
- Rutledge, R. B., Lazzaro, S. C., Lau, B., Myers, C. E., Gluck, M. A., & Glimcher, P. W. (2009). Dopaminergic drugs modulate learning rates and perseveration in Parkinson's patients in a dynamic foraging task. *Journal of Neuroscience*, 29(48), 15104–15114
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306), 1593–1599
- Schutte, I., Slagter, H. A., Collins, A. G., Frank, M. J., & Kenemans, J. L. (2017). Stimulus discriminability may bias value-based probabilistic learning. *PLoS one*, 12(5), e0176205
- Sense, F., Behrens, F., Meijer, R. R., & van Rijn, H. (2016). An individual's rate of forgetting is stable over time but differs across materials. *Topics in cognitive science*, 8(1), 305–321
- Sense, F., Meijer, R. R., & van Rijn, H. (2018). Exploration of the Rate of Forgetting as a Domain-Specific Individual Differences Measure. In *Frontiers in Education* (Vol. 3, p. 112)
- Shohamy, D., Myers, C. E., Kalanithi, J., & Gluck, M. A. (2008). Basal ganglia and dopamine contributions to probabilistic category learning. *Neuroscience & Biobehavioral Reviews*, 32(2), 219–236
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin*, 86(2), 420–428
- Sian, J., Youdim, M. B. H., Riederer, P., & Gerlach, M. (1999). Biochemical anatomy of the basal ganglia and associated neural systems. In *Basic Neurochemistry: Molecular, Cellular and Medical Aspects* (6th ed.). Philadelphia: Lippincott-Raven
- Simon, J. R., Howard Jr., J. H., & Howard, D. V. (2010). Adult age differences in learning from positive and negative probabilistic feedback. *Neuropsychology*, 24(4), 534–541
- Stocco, A. (2018). A biologically-plausible action selection system for cognitive architectures: Implications of basal ganglia anatomy for learning and decision making models. *Cognitive Science*, 42(2018), 457–490
- Stocco, A., & Anderson, J. R. (2008). Endogenous control and task representation: an fMRI study in algebraic problem-solving. *Journal of Cognitive Neuroscience*, 20(7), 1300–1314.
- Stocco, A., Murray, N. L., Yamasaki, B. L., Renno, T. J., Nguyen, J., & Prat, C. S. (2017). Individual differences in the Simon effect are underpinned by differences in the competitive dynamics in the basal ganglia: An experimental verification and a computational model. *Cognition*, 164, 31–45
- Stocco, A., Prat, C. S., & Graham, L. K. (in press). Individual differences in reward-based learning predict fluid reasoning abilities. *Cognitive Science*
- Taylor, P. J. (2010). An introduction to intraclass correlation that resolves some common confusions. Unpublished manuscript, University of Massachusetts, Boston, USA. Retrieved from http://www.faculty.umb.edu/peter_taylor/09b.Pdf
- Voon, V., Pessiglione, M., Brezing, C., Gallea, C., Fernandez, H. H., Dolan, R. J., & Hallett, M. (2010). Mechanisms underlying dopamine-mediated reward bias in compulsive behaviors. *Neuron*, 65(1), 135–142
- Waltz, J. A., Frank, M. J., Robinson, B. M., & Gold, J. M. (2007). Selective reinforcement learning deficits in schizophrenia support predictions from computational models of striatal-cortical dysfunction. *Biological psychiatry*, 62(7), 756–764
- Weismueller, B., Ghio, M., Logmin, K., Hartmann, C., Schnitzler, A., Pollok, B., et al. (2018). Effects of feedback delay on learning from positive and negative feedback in patients with Parkinson's disease off medication. *Neuropsychologia*, 117, 46–54
- White, C. N., Congdon, E., Mumford, J. A., Karlsgodt, K. H., Sabb, F. W., Freimer, N. B., London, E. D., Cannon, T. D., Bilder, R. M., & Poldrack, R. A. (2014). Decomposing decision components in the stop-signal task: a model-based approach to individual differences in inhibitory control. *Journal of cognitive neuroscience*, 26(8), 1601–1614
- White, C. N., Curl, R. A., & Sloane, J. F. (2016). Using decision models to enhance investigations of individual differences in cognitive neuroscience. *Frontiers in Psychology*, 7, 81
- Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience*, 7(6), 464–476.
- Zhang, Y., Bertolino, A., Fazio, L., Blasi, G., Rampino, A., Romano, R., et al. (2007). Polymorphisms in human dopamine D2 receptor gene affect gene expression, splicing, and neuronal activity during working memory. *Proceedings of the National Academy of Sciences*, 104(51), 20552–20557

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.