Test-retest reliability of task-based measures of voluntary persistence

Yixin Chen (yxyxchen@bu.edu)
Department of Psychological and Brain Sciences, Boston University, MA 02215 USA

Daniel Fulford(dfulford@bu.edu)
Sargent College of Health and Rehabilitation Sciences, Boston University, MA 02215 USA
Department of Psychological and Brain Sciences, Boston University, MA 02215 USA

Joseph T. McGuire (jtmcg@bu.edu)
Department of Psychological and Brain Sciences, Boston University, MA 02215 USA
Center for Systems Neuroscience, Boston University, MA 02215 USA

Abstract

Decision makers face a nontrivial problem when evaluating how much time to invest in an uncertain future prospect. Unconditional persistence is not always advantageous; rather, different levels of persistence are favored in environments with different temporal statistics. Previous studies using foraging-like decision-making tasks have found that people can rapidly recalibrate their persistence behavior—becoming either more or less willing to tolerate delay—after a short period of direct experience with the temporal statistics of a new environment. Furthermore, substantial individual variation is apparent both in baseline levels of persistence and in the flexibility of recalibration across environments. However, it is unknown to what degree such variation reflects trait-like individual differences in contrast to session-specific measurement noise. Here we investigated the test-retest reliability of individual variation in behavioral persistence in a computerized decision-making task. We conducted an online experiment in which participants (n=141 after exclusions) performed the task on two occasions separated by a three-week interval. We evaluated the test-retest reliability of several behavior-derived indices, including: a descriptive estimate of overall willingness to wait, a contrast measure reflecting flexibility of recalibration across environments, and individual-level parameter estimates derived from a reinforcement learning model of adaptive persistence. The results showed strong evidence for stable, trait-like individual variation in multiple aspects of persistence-related decision-making behavior. Our findings establish a foundation for future investigations of associations between task-derived parameters of decision behavior and other cognitive and motivational traits.

Keywords: decision making; intertemporal choice; foraging; test-retest reliability; computational modeling; reinforcement learning

Introduction

Although the ability to delay gratification is essential for attaining beneficial long-run outcomes, not all outcomes in the real world are worth waiting for indefinitely. A general problem faced by both humans and foraging animals in delay-of-gratification scenarios is to regulate persistence in a context-appropriate manner (McNamara, 1982; McGuire & Kable, 2013; Fawcett, McNamara, & Houston, 2012). The cognitive processes that govern the contextually adaptive regulation of persistence—and the ways in which such cognitive processes might differ across individuals—are not yet fully understood.

One behavioral paradigm for studying the experience-driven calibration of persistence is the *willingness-to-wait* task (McGuire & Kable, 2015, 2012; Lempert, McGuire, Hazeltine, Phelps, & Kable, 2018). In this task, participants continuously decide whether (and for how long) to continue waiting for delayed and temporally uncertain rewards. The alternative to waiting is to disengage from the current reward opportunity and move on to a new one. Akin to foraging paradigms (McNamara, 1982; Constantino & Daw, 2015), participants' goal in the task is to maximize the *rate* of reward accrual over the course of a fixed time period by choosing between exploiting and abandoning individually encountered reward prospects. The probability distributions governing delay durations are manipulated across environments so that either higher or lower levels of persistence lead to more advantageous outcomes. In *high persistence* (HP) environments, the passage of time supports an inference that the reward's expected arrival time is drawing nearer, whereas in *limited-persistence* (LP) environments, time passage beyond a certain point signals that the expected remaining delay is growing longer. Using this paradigm, researchers have consistently observed an overall trend toward adaptive calibration (greater willingness to wait in HP environments than in LP environments) accompanied by an appreciable level of variation across individuals.

Laboratory measures of intertemporal decision making can, in principle, offer insight about sources of variation in motivational traits such as impulsivity that are relevant to mental health and well being. For example, estimates of temporal discount rates show strong test-retest reliability (Kirby & Kirby, 2009) and are correlated with college GPA (Kirby, Winston, Santiesteban, & Kirby, 2005), the incidence of substance use disorder (Kirby, Petry, & Bickel, 1999), and problematic drinking behavior (Vuchinich & Simpson, 1998). A laboratory measure of delay of gratification in children (the "marshmallow test") has been found to correlate with intelligence, social skills, cognitive control, academic achievement and other competencies years later (Mischel, Shoda, & Rodriguez, 1989; Casey et al., 2011; Mischel, Shoda, & Peake, 1988). However, task-derived measures generally exhibit lower reliability compared to self-report measures, raising doubts about their general utility for investigating individual differences (Hedge, Powell, & Sumner, 2017; Enkavi et al., 2019).

The present work assessed the test-retest reliability of individual differences in behavior in the willingness-to-wait task.

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

Such differences can be operationalized in terms of several different behavior-derived indices. One possibility is to calculate a descriptive statistic for the average amount of time an individual was willing to wait, conditional on the reward not having been delivered. A second possibility is to calculate the difference between waiting times in HP and LP environments, with a larger difference presumably indicating greater adaptive flexibility. This type of *contrast measure* can more specifically isolate a putative cognitive factor (e.g., flexibility); however, contrast measures are believed to be less reliable than raw task measures because subtraction of random variables mathematically increases their variance (Caruso, 2004).

A third possibility is to fit a behavior-generating computational model to the task data and capture individual differences in model parameter estimates. Individually fit model parameters are believed to provide deeper insights into cognitive and biological processes than conventional task measures (Maia & Frank, 2011; Patzelt, Hartley, & Gershman, 2018; Gläscher, Adolphs, & Tranel, 2019; Frank, Moustafa, Haughey, Curran, & Hutchison, 2007). However, previous results on the reliability of such estimates are mixed. One study reported certain parameters of diffusion models were as stable as raw task measures (e.g., drift rate, threshold and nondecision time). However, another study found that parameters of reinforcement learning (RL) models showed either weak or no correlations over repeated administration (Weidinger, Gradassi, Molleman, & van den Bos, 2019).

We conducted a large-sample test-retest study online to assess the reliability of three types of measures in the willingness-to-wait task: raw task measures, contrast measures, and model-derived parameters.

Methods

Participants

The test-retest study was conducted on Amazon Mechanical Turk. Participants completed a 20-minute willingness-to-wait task in two sessions spaced approximately three weeks apart. 197 participants participated in the first session. 171 participants whose data met quality criteria were invited to participate in the second session (see **Data exclusion**). 141 participants (37.5% female, age: mean = 42.2, median = 40.0, range = 23–74) satisfactorily completed both sessions.

Task

For this online study, we used an active version of the willingness-to-wait task to reduce the risk of attentional lapses. Unlike earlier versions of the task (McGuire & Kable, 2015, 2012), participants had to make repeated key-presses throughout the delay interval in order to continue waiting.

Figure 1a shows a schematic of the sequence of task events. At the onset of a trial, a token worth 0¢ appeared in the center of the screen, with the white text 'start pressing now' above. Once the participant started to make key presses, the white text disappeared and the token would mature to a value of 2¢ after a random delay. Participants could either make continuous key presses (≥ 2 strokes per second) to keep waiting for the token or stop pressing anytime to sell the token. Once the token was sold, the word 'SOLD' appeared in red over the token for 1 s. After a 0.5-s blank screen, a new trial started. The interface also displayed the accumulated earnings and the time remaining in the current block at the bottom of the screen.

In each session, the participant experienced a 10-min block in the LP environment followed by a 10-min block in the HP environment. As shown in Figure 1b, in the HP environment delays were sampled equally from eight possible values evenly spaced from 1.5 to 12 s, whereas in the LP environment delays were equally sampled from eight values logarithmically spaced up to 24 s. In the HP environment the optimal strategy was always to wait, whereas in the LP environment the optimal strategy was to wait up to 2.3 s (see Normative analysis for details).

A fixed ordering of the two timing condition conditions (LP followed by HP in each session) was used to avoid introducing extraneous variance across participants and sessions. This design feature reflects that the study's main goals focused on individual differences rather than on main effects of condition (which have been demonstrated previously using between-subject designs or counterbalanced condition orders).

Different environments were represented by different token colors. The color mapping differed across sessions and was counterbalanced across participants. For example, if the LP environment was represented by purple tokens in the first session, it would be presented by pink tokens in the second session, making the environment visually novel. This design feature was intended to mitigate practice effects.

In each session, participants completed four practice trials before the actual task. In the first two trials, they were required to wait until the token matured. In the next two trials, they were required to sell the token before it matured. No environment-specific information was provided in the practice.

Data exclusion

Participants were excluded if they failed to complete the entire session, if they sold matured tokens too slowly (median RT ≥ 1.2 s), or if they achieved ≤ 450 s of active on-task time in a block (because of cumulative delay in starting new trials or selling matured tokens).

Since participants showed a tendency to transiently decrease their persistence level under time pressure near the end of a block, we excluded responses in the last 24 s of each block in all of our analyses except for generating the entire learning course in Figure 2a.

Behavioral analysis

To characterize behavior over time, we calculated a local estimate of each participant's willingness to wait (WTW) every 2 s throughout the experiment. During quit trials, this estimate

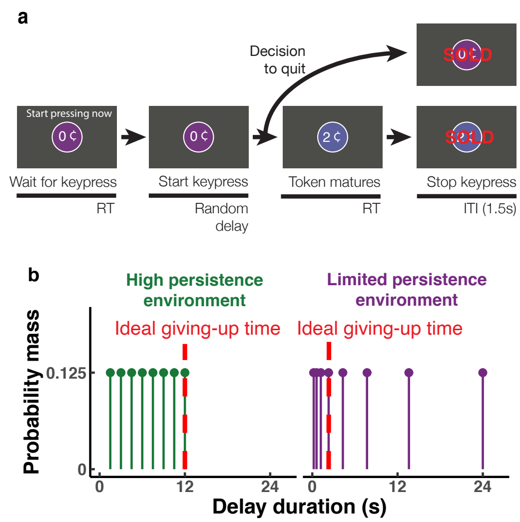

Figure 1: Sequence of events in the willingness-to-wait task (a) and delay durations that defined the high-persistence and limited-persistence environments (b).

was set equal to the observed waiting time. During other trials, the estimate was set equal to the longest time waited since the last quit trial. The estimate was capped at 12 s to make the two environments comparable.

To summarize each individual's overall level of behavioral persistence, we constructed a Kaplan-Meier survival curve for each participant in each block. The analysis was restricted to the 0 to 12 s range for which we had observations in both environments. Rewarded trials were treated as censored observations of the participant's waiting time. The area under the survival curve (AUC) captured the average amount of time a participant was willing to wait within the first 12 s, conditional on the reward not having been delivered. The same analysis was also performed on the first and second half of each block separately to summarize behavior at a finer scale.

To assess the flexibility with which participants recalibrated persistence across environments, we calculated ΔAUC by subtracting AUC in the LP environment from that in the HP environment. A larger value of ΔAUC was interpreted as reflecting greater flexibility.

Normative analysis

We determined the reward-maximizing behavioral strategy in each environment by calculating the average rate of return for various waiting policies. Define the length of delay for a token to mature as a random variable, λ. For a policy of quitting at time T, define the average rate of return as ρ_T, and the probability that the token matures before T as $P_T = P(\lambda \leq T)$. Define W_T as the expected delay if the reward is received before T, $W_T = E[\lambda \mid \lambda \leq T]$. The expected rate of return, in ¢/s, is

$$\rho_T = \frac{2 \cdot P_T}{P_T W_T + (1 - P_T)T + 1.5}$$

The numerator is a trial's expected gain in cents and the denominator is a trial's expected duration in seconds (including the 1.5-s ITI). The optimal giving-up time is simply the value of T that maximizes ρ_T, denoted as T^*. The corresponding expected rate of return is denoted as ρ^*_T.

In our task, the optimal giving-up time was 12 s in the HP environment ($\rho^*_T = 0.242$ ¢/s), and 2.312 s in the LP environment ($\rho^*_T = 0.313$ ¢/s)

Modeling analysis

Details of the RL model for the willingness-to-wait task can be found elsewhere (work in preparation). Briefly, the model conceptualizes behavior in the task as emerging from a series of wait-or-quit choices over the course of the delay. Each time step in a trial is a distinct temporal state. Decision makers assess the value of continuing to wait in comparison to quitting for each specific temporal state and make a choice accordingly.

The RL model has five parameters, each of which reflects a specific computational factor. The parameters can take individual-specific values to capture variation in behavior: learning rate α, valence-dependent bias ν, inverse temperature of the decision function τ, temporal discounting γ, and prior beliefs about the value of waiting η.

Model parameters were estimated using Bayesian methods implemented in Stan (Carpenter et al., 2017), with uniform priors: $\alpha \sim \text{unif}(0, 0.3)$, $\nu \sim \text{unif}(0, 5)$, $\tau \sim \text{unif}(0.1, 22)$, $\gamma \sim \text{unif}(0.7, 1)$, $\eta \sim \text{unif}(0, 6.5)$. We fit each model for each participant and each session with 4 chains, 4000 samples per chain. The first 2000 samples in each chain were discarded as burn-in. The mean of the posterior distribution for each parameter was used as a point estimate.

We compared the full version of the model (described above) with a reduced version that lacked the valence-dependent bias parameter ν and therefore had equal sensitivity to rewarding and nonrewarding outcomes. We chose *widely applicable information criterion* (WAIC) as our model fit criterion, a quantity that rewards goodness of fit while penalizing model complexity. To determine which model provided the better fit at the group level, we tabulated the number of participants best fit by each form of the model.

Reliability analysis

We used Spearman's ρ as our measure of test-retest reliability. 95% confidence intervals for Spearman's ρ were constructed using bootstrap methods (1000 samples). The magnitudes of two Spearman's ρ coefficients were compared using permutation tests (1000 permutations, Omelka & Pauly, 2012).

Results

Behavioral results

Participants were able to calibrate their persistence level in a context-dependent manner. Figure 2a shows the local WTW estimate as a function of task time. In both sessions, participants reduced their persistence level in the LP environment and increased their persistence level in the HP environment. We then used survival analysis to estimate each participant's overall probability of 'surviving' various lengths of time without quitting. Figure 2b shows averaged per-participant empirical survival curves in the two environments. The area under the curve (AUC) estimates how much of the first 12 s a participant was willing to wait on average. In both sessions, AUC values were higher in the HP environment than in the LP environment (Session 1: HP, median = 7.52 s, IQR = 5.41-9.21 s; LP, median = 5.88 s, IQR = 4.73-7.73 s, signed-rank $p < 0.001$. Session 2: HP, median = 6.86 s, IQR = 5.36-9.08 s; LP, median = 5.26, IQR = 4.04-6.53 s, signed-rank $p < 0.001$).

There was a concern that behavior might systematically differ across sessions due to experience with the task, which might have unpredictable effects on estimates of test-retest reliability. To access the severity of this issue, we first compared participants' local behavior across sessions (Figure 2a). We found the local WTW estimate at the beginning of the task was significantly lower in Session 2 compared to Session 1. This could have occurred because participants updated their prior belief about the general times scale of rewards in the task. However, no systematic divergence was observed within other time windows, either in terms of persistence level or the calibration trajectory. We then calculated AUC values for the first and second halves of each block and compared them across sessions. Consistent with the local comparison results, significant differences in AUC were only detected in the first half of the LP block (Session 1: median = 6.66 s, IQR = 5.29-8.00 s; Session 2: median = 5.31 s, IQR = 4.24-7.07; signed-rank $p < 0.001$).

Modeling results

Participants were excluded in modeling analyses if the MCMC chains fit to their data did not converge. For the full model, seven participants in Session 1 and three participants in Session 2 were excluded. For the reduced model, seven participants in Session 1 and eight participants in Session 2 were excluded.

Consistent with previous findings (work in preparation), the full model with two separate learning rates for rewarding and nonrewarding outcomes fit the data better than the reduced model. It had lower WAIC and better fit the majority of participants in both sessions (Table 1). Table 2 summarizes the individually fit parameters for the full model.

To validate the parameter estimates of the full model, we simulated 10 data sets with each participant's estimated parameter values and compared the AUC value calculated from the model-generated data (averaged across 10 data sets) with the observed AUC for that participant. As shown in Figure 3, with individually fit parameters, the full RL model was highly accurate in reproducing the data.

Reliability of descriptive task measures

AUC was the main descriptive task measure in the willingness-to-wait task. We first assessed the reliability of

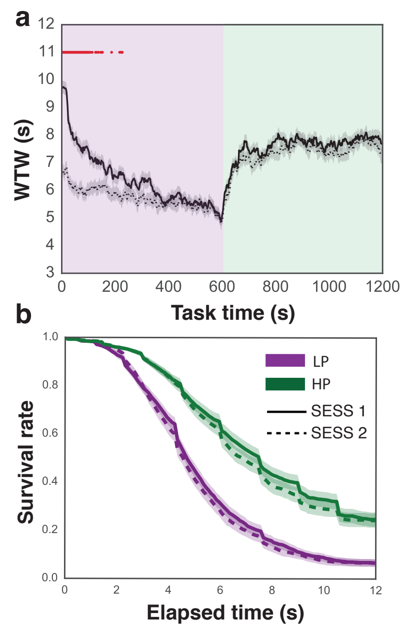

Figure 2: Behavioral results. (a) Mean local WTW estimate averaged across participants (with standard error of the mean; s.e.m.), sampled at intervals of 2 s. Red dots at the top indicate time points at which the local WTW estimate was significantly different between Sessions 1 and 2 (signed-rank $p < 0.05$ after false discovery rate correction). (b) Mean survival curves averaged across participants, with s.e.m., sampled at intervals of 0.1 s.

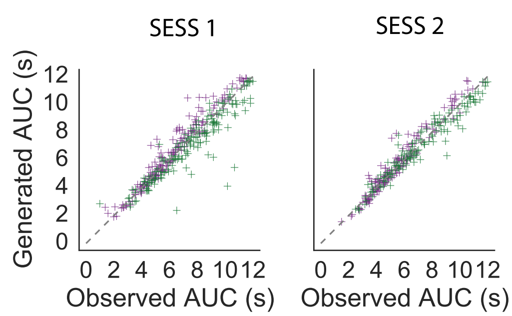

Figure 3: Observed AUC compared with AUC generated by the full RL model. Each data point is the AUC value of one participant in one environment. The diagonal line is plotted for reference.

Table 1: Model comparison results

		Reduced model	Full model
WAIC	Session 1	393.24 ± 11.60	382.93 ± 10.82
(mean ± s.e.m.)	Session 2	374.07 ± 10.95	367.89 ± 10.68
Best explained	Session 1	36	112
	Session 2	48	83

Table 2: Parameter estimates of the full RL model. Data are reported as median (IQR).

Free parameter	Session 1	Session 2
Learning rate α	0.010 (0.004-0.023)	0.005 (0.002-0.017)
Valence-dependent bias ν	0.490 (0.160-1.020)	0.712 (0.345-1.524)
Inverse temperature τ	5.903 (4.051-8.069)	7.508 (5.179-11.083)
Temporal discounting γ	0.858 (0.827-0.897)	0.831 (0.797-0.866)
Prior belief parameter η	0.896 (0.660-1.249)	0.639 (0.448-0.910)

the AUC calculated based on data from each block (HP and LP). As shown in Figure 4, the reliability of AUC was 0.570 (95% CI [0.423, 0.696]) in the LP environment and 0.737 (95% CI [0.659, 0.806]) in HP environment.

We then assessed the reliability of the AUC calculated based on data from each half block (LP-1st, LP-2nd, HP-1st, and HP-2nd). The goal was to examine whether the reliability of AUC was time-dependent. AUC in the first half of the LP block was the least reliable (LP-1st = 0.420, 95% CI [0.251, 0.546], LP-2nd = 0.619, 95% CI [0.481, 0.732], HP-1st = 0.683, 95% CI [0.578, 0.766], HP-2nd = 0.683, 95% CI [0.573, 0.769]). This was probably because this time period was most affected by practice effects as discussed above.

Reliability of contrast measures

ΔAUC was calculated by subtracting AUC in the LP environment from that in the HP environment. The resulting score can be interpreted as reflecting an individual's flexibility in calibrating persistence across environments. Compared to AUC, ΔAUC had lower reliability (Spearman's ρ = 0.492, 95% CI [0.351, 0.604], Figure 5).

Reliability of RL model parameters

Figure 6 compares the parameter estimates of the full RL model across sessions. All model parameters were significantly correlated across sessions. However, their reliability was relatively low compared to AUC and ΔAUC (Spearman's ρ: α = 0.344, 95% CI [0.192, 0.481]; ν = 0.275, 95% CI [0.12, 0.422]; τ = 0.353, 95% CI [0.196, 0.487]; γ = 0.191, 95% CI [0.03, 0.346]; η = 0.316, 95% CI [0.145, 0.464]).

Comparing the full model and the reduced model (Figure 7), we observed a nonsignificant tendency for the full model to yield a more reliable estimate of the learning rate α (Spearman's ρ: full model = 0.344, 95% CI [0.192, 0.481]; reduced model = 0.196, 95% CI [0.006, 0.351]). Such an effect could potentially have occurred because the full model was able to account for learning asymmetry for rewarding and non-rewarding outcomes with the additional parameter, valence-dependent bias ν. The reliability of other model parameters was roughly the same.

Conclusion

Maladaptive persistence behavior is closely associated with impulsivity, a transdiagnostic feature of multiple mental health conditions. Persistence calibration is potentially a

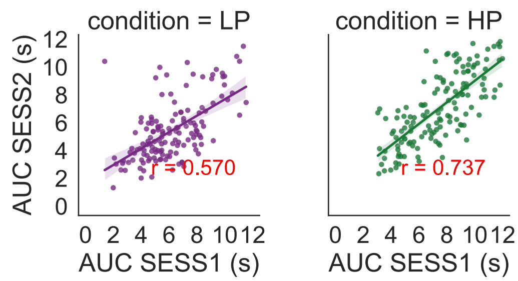

Figure 4: Across-session correlations of AUC estimates. Spearman's correlation statistics are annotated.

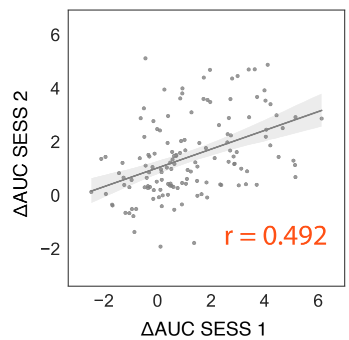

Figure 5: Across-session correlation of ΔAUC. For illustrative purposes, outlier data points (> 1.5 IQR, n = 3) are not shown in the scatterplot. Spearman's correlation statistics were calculated using all data points.

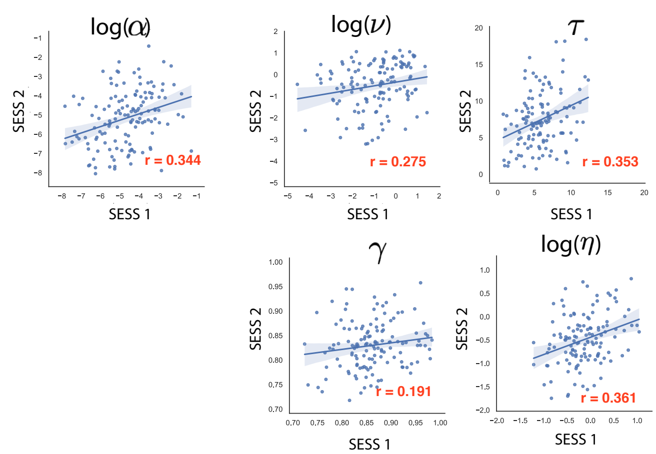

Figure 6: Across-session correlations of individually estimated parameters for the full RL model. For illustrative purposes, three parameters with heavy-tailed distributions were log transformed and outlier data points (> 1.5 IQR) are not shown in scatterplots. Spearman's correlation statistics were calculated using all data points.

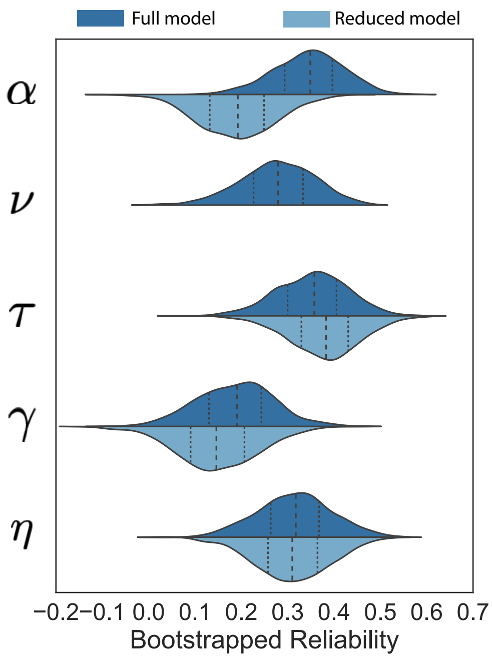

Figure 7: Violin plots show bootstrapped Spearman's ρ estimates for model parameters of the full model and the reduced model.

promising factor to test for associations with clinical symptom dimensions. However, an important prerequisite for such investigations is to assess the reliability with which individual-level parameters of behavioral persistence can be measured in laboratory settings. Previous studies have developed a behavioral task that examines context-dependent persistence under temporal uncertainty, but the test-retest reliability of task behavior has yet to be assessed.

In this study, an online test-retest data set was collected using the willingness-to-wait task, with a final sample size of 141 and an inter-test interval of 3 weeks. We assessed the test-retest reliability of three types of measures: raw task measures (AUC), contrast measures (ΔAUC), and RL model parameters.

All task measures were significantly correlated across sessions and captured substantial individual differences. The raw task measure, AUC, reached an average Spearman's ρ of 0.654 across environments, which was comparable to measures obtained using standard self-regulation tasks (Enkavi et al., 2019). This provided preliminary evidence that persistence calibration is a stable trait and qualifies for further investigation as a predictor of real-life outcomes.

In considering the relative reliability of different types of task-derived indices, our study found qualitatively similar results to earlier studies (Enkavi et al., 2019; Weidinger et al., 2019). Raw task measures, AUC measured in the two environments, were the most reliable, with Spearman's ρ = 0.570 in the LP environment and Spearman's ρ = 0.737 in the HP environment. In comparison, the contrast measure, ΔAUC, was less reliable (Spearman's ρ = 0.492). The RL model parameters were significantly correlated across sessions but were the least reliable (average Spearman's ρ = 0.296).

Test-retest reliability was also influenced by other methodological factors. For example, we found the better-fit variant of the RL model showed a tendency toward yielding more reliable learning rate estimates. This underscores the point that good alignment of the model and the paradigm is crucial for obtaining reliable parameter estimates. To improve the reliability of model parameters, it will be important to examine the effects of methodological factors (e.g., choice of computational models, fitting procedures, and data prepossessing methods) more comprehensively in future work.

References

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1). doi: 10.18637/jss.v076.i01

Caruso, J. C. (2004). A comparison of the reliabilities of four types of difference scores for five cognitive assessment batteries. *European Journal of Psychological Assessment*, *20*(3), 166–171. doi: 10.1027/1015-5759.20.3.166

Casey, B. J., Somerville, L. H., Gotlib, I. H., Ayduk, O., Franklin, N. T., Askren, M. K., ... Shoda, Y. (2011). Behavioral and neural correlates of delay of gratification 40 years later. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(36), 14998–15003. doi: 10.1073/pnas.1108561108

Constantino, S. M., & Daw, N. D. (2015). Learning the opportunity cost of time in a patch-foraging task. , *15*(4), 837–853. doi: 10.3758/s13415-015-0350-y

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 of the United States of America*. doi: 10.1073/pnas.1818430116

Fawcett, T. W., McNamara, J. M., & Houston, A. I. (2012). When is it adaptive to be patient? A general framework for evaluating delayed rewards. *Behavioural Processes*, *89*(2), 128–136. doi: 10.1016/j.beproc.2011.08.015

Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchison, K. E. (2007). 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. doi: 10.1073/pnas.0706111104

Gläscher, J., Adolphs, R., & Tranel, D. (2019). Model-based lesion mapping of cognitive control using the Wisconsin Card Sorting Test. *Nature communications*, *10*(1), 20. doi: 10.1038/s41467-018-07912-5

Hedge, C., Powell, G., & Sumner, P. (2017). The reliability paradox: Why robust cognitive tasks do not produce reli-

able individual differences.
doi: 10.3758/s13428-017-0935-1

Kirby, K. N., & Kirby, K. N. (2009). One-year temporal stability of delay-discount rates A-2. *Psychonomic Bulletin Review*, *16*(3), 457–462. doi: 10.3758/PBR.16.3.457

Kirby, K. N., Petry, N. M., & Bickel, W. K. (1999). Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *Journal of Experimental Psychology: General*. doi: 10.1037/0096-3445.128.1.78

Kirby, K. N., Winston, G. C., Santiesteban, M., & Kirby, . K. N. (2005). Impatience and grades: Delay-discount rates correlate negatively with college GPA. *Learning and Individual Differences*, *15*, 213–222. Retrieved from `www.elsevier.com/locate/lindif` doi: 10.1016/j.lindif.2005.01.003

Lempert, K. M., McGuire, J. T., Hazeltine, D. B., Phelps, E. A., & Kable, J. W. (2018, feb). The effects of acute stress on the calibration of persistence. *Neurobiology of Stress*, *8*, 1–9. Retrieved from `https://doi.org/10.1016/j.ynstr.2017.11.001` doi: 10.1016/j.ynstr.2017.11.001

Maia, T. V., & Frank, M. J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nature Neuroscience*, *14*(2), 154–162. doi: 10.1038/nn.2723

McGuire, J. T., & Kable, J. W. (2012). Decision makers calibrate behavioral persistence on the basis of time-interval experience.
doi: 10.1016/j.cognition.2012.03.008

McGuire, J. T., & Kable, J. W. (2013). Rational temporal predictions can underlie apparent failures to delay gratification. *Psychological Review*, *120*(2), 395–410. doi: 10.1037/a0031910

McGuire, J. T., & Kable, J. W. (2015). Medial prefrontal cortical activity reflects dynamic re-evaluation during voluntary persistence. *Nature Neuroscience*, *18*(5), 760–766. doi: 10.1038/nn.3994

McNamara, J. (1982, apr). Optimal patch use in a stochastic environment. *Theoretical Population Biology*, *21*(2), 269–288. doi: 10.1016/0040-5809(82)90018-1

Mischel, W., Shoda, Y., & Peake, P. K. (1988). The Nature of Adolescent Competencies Predicted by Preschool Delay of Gratification. *Journal of Personality and Social Psychology*, *54*(4), 687–696. doi: 10.1037/0022-3514.54.4.687

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science*. doi: 10.1126/science.2658056

Omelka, M., & Pauly, M. (2012). Testing equality of correlation coefficients in two populations via permutation methods. *Journal of Statistical Planning and Inference*, *142*(6), 1396–1406. doi: 10.1016/j.jspi.2011.12.018

Patzelt, E. H., Hartley, C. A., & Gershman, S. J. (2018). Computational Phenotyping: Using Models to Understand Individual Differences in Personality, Development, and Mental Illness. *Personality Neuroscience*, *1*. Retrieved from `https://doi.org/10.1017/pen.2018.14` doi: 10.1017/pen.2018.14

Vuchinich, R. E., & Simpson, C. A. (1998). Hyperbolic temporal discounting in social drinkers and problem drinkers. *Experimental and Clinical Psychopharmacology*, *6*(3), 292–305. Retrieved from `/record/1998-04990-007` doi: 10.1037/1064-1297.6.3.292

Weidinger, L., Gradassi, A., Molleman, L., & van den Bos, W. (2019). Test-retest reliability of canonical reinforcement learning models.. doi: 10.32470/ccn.2019.1053-0