

1 Reliability of Decision-Making and Reinforcement Learning

2 Computational Parameters

3 Anahit Mkrtchian^{1*}, Vincent Valton¹, Jonathan P. Roiser¹

4

5 ¹Neuroscience and Mental Health Group, Institute of Cognitive Neuroscience, University
6 College London, London, United Kingdom.

7

8 *Correspondence

9 Anahit Mkrtchian
10 Institute of Cognitive Neuroscience, University College London
11 Alexandra House, 17-19 Queen Square
12 London, WC1N 3AZ
13 United Kingdom
14 Email: a.mkrtchian@ucl.ac.uk

15

16 **Running Title:** Reliability of computational parameters

17

18 **Keywords:** Reliability, Computational psychiatry, Decision-making, Reinforcement learning,
19 Prospect theory, Gambling

1 Abstract

2 Computational models can offer mechanistic insight into cognition and therefore have the
3 potential to transform our understanding of psychiatric disorders and their treatment. For
4 translational efforts to be successful, it is imperative that computational measures capture
5 individual characteristics reliably. To date, this issue has received little consideration. Here we
6 examine the reliability of reinforcement learning and economic models derived from two
7 commonly used tasks. Healthy individuals (N=50) completed a restless four-armed bandit and a
8 calibrated gambling task twice, two weeks apart. Reward and punishment learning rates from
9 the reinforcement learning model showed good reliability and reward and punishment
10 sensitivity from the same model had fair reliability; while risk aversion and loss aversion
11 parameters from a prospect theory model exhibited good and excellent reliability, respectively.
12 Both models were further able to predict future behaviour above chance within individuals.
13 This prediction was better when based on participants' own model parameters than other
14 participants' parameter estimates. These results suggest that reinforcement learning, and
15 particularly prospect theory parameters, as derived from a restless four-armed bandit and a
16 calibrated gambling task, can be measured reliably to assess learning and decision-making
17 mechanisms, and that these processes may represent relatively distinct computational profiles
18 across individuals. Overall, these findings indicate the translational potential of clinically-
19 relevant computational parameters for precision psychiatry.

20

21 Keywords: Computational psychiatry, Reliability, Decision-making, Reinforcement Learning,
22 Prospect Theory, Gambling

1 Introduction

2 Cognitive and neural processes are increasingly conceptualized in computational terms

3 (Palminteri et al., 2017). Generative computational models offer the advantage of examining

4 behaviourally unobservable, but important, latent processes that drive behaviour and can be

5 closely linked to neurobiology (Huys et al., 2016; Montague et al., 2004). As such, they provide

6 a mathematically precise framework for specifying hypotheses about the cognitive processes

7 that generate behaviour. These features make modelling a powerful tool to provide

8 mechanistic accounts into (a)typical behaviours, including those associated with psychiatric

9 symptoms.

10

11 The increasing adoption of computational approaches in cognitive neuroscience inspired the

12 emerging discipline of computational psychiatry, which aims to better understand mental

13 illness through computational methods with the ultimate goal of transforming such knowledge

14 into new personalised treatment strategies (Adams et al., 2016; Browning et al., 2020; Huys,

15 2018; Huys et al., 2016; Huys et al., 2011; Maia & Frank, 2011; Montague et al., 2012; Patzelt et

16 al., 2018; Paulus et al., 2016; Paulus & Thompson, 2019; Teufel & Fletcher, 2016; Wang &

17 Krystal, 2014; Wiecki et al., 2015). For such translational endeavours to be successful, however,

18 it is vital that computational measures capture individual characteristics reliably (Browning et

19 al., 2020; Paulus et al., 2016). Specifically, the reliability of measures set an upper limit for

20 detecting both relationships with other measures, e.g., symptoms, and the effect of

21 interventions in e.g., randomized controlled trials, which may otherwise be obscured by poor

22 test-retest reliability.

1
2 One area of computational modelling that has received particular attention is the reward and
3 punishment processes underlying decision-making (Maia, 2009). In particular, two classes of
4 computational models, respectively originating from computer science and behavioural
5 economics, have been influential in characterizing the cognitive mechanisms underlying
6 decision-making: reinforcement learning (RL) and prospect theory (PT) (Dayan & Niv, 2008;
7 Kahneman & Tversky, 1979; Maia, 2009; Niv, 2009; Schonberg et al., 2011; Sutton & Barto,
8 2018; Tversky & Kahneman, 1992). RL models describe how agents learn from rewards and
9 punishments through trial-and-error (Sutton & Barto, 2018). Within the field of computational
10 psychiatry, this set of models has perhaps been the most influential. For example, reward and
11 punishment sensitivity (reflecting subjective valuation of the outcomes) and learning rate
12 (reflecting how quickly individuals learn from better- or worse-than-expected outcomes) have
13 been associated with distinct symptomatology and neural signals (Daw & Doya, 2006; Huys et
14 al., 2021; Maia & Frank, 2011; Niv, 2009). A commonly-used RL task is the multi-armed
15 “bandit”, in which individuals choose between multiple slot machines with fluctuating,
16 unknown probabilities of reward and punishment, with the goal of maximizing earnings (Daw et
17 al., 2006; Seymour et al., 2012; Speekenbrink & Konstantinidis, 2015; Yi et al., 2009).
18 Participants must decide on each trial whether to persist with the previously sampled slot
19 machine or explore others which may yield better outcomes. The mechanisms thought to
20 underlie these decisions can be captured effectively by RL models (Aylward et al., 2019; Daw et
21 al., 2006).
22

1 PT models, on the other hand, describe the cognitive processes driving decision-making biases
2 under known risks, and have been extremely influential in understanding economic decision-
3 making (Kahneman & Tversky, 1979; Ruggeri et al., 2020; Schonberg et al., 2011; Sokol-Hessner
4 & Rutledge, 2019; Tversky & Kahneman, 1992). These processes are often examined by asking
5 participants to choose between a guaranteed outcome (e.g., £0 gain) and a 50% gamble with
6 two possible outcomes (e.g., £30 gain or £10 loss). It is commonly observed that humans tend
7 to prefer a sure payment over a risky payment with equivalent or higher expected value. For
8 example, you may prefer an investment with a fixed return over one with a potentially higher
9 but uncertain return. PT proposes that these observations can be accounted for by two
10 different processes: 1) risk aversion – the preference for certain over uncertain gains, and 2)
11 loss aversion – weighting losses more heavily than gains. Risk and loss aversion vary across
12 individuals, and these differences have been associated with various psychiatric states, such as
13 greater risk aversion in anxiety, increased loss aversion in obsessive-compulsive disorder and
14 associations between loss aversion and suicidality (Baek et al., 2017; Brown et al., 2013;
15 Charpentier et al., 2017; Charpentier et al., 2016; Hadlaczky et al., 2018; Hartley & Phelps,
16 2012; Klaus et al., 2020; Sip et al., 2017; Stauffer et al., 2014; Tobler et al., 2009; Tremeau et al.,
17 2008). Importantly, computational modelling has allowed researchers to dissociate risk and loss
18 aversion and their contribution to symptoms (Charpentier et al., 2017).

19
20 Parameters from RL and PT models show promise in generating insights into the mechanisms
21 underlying psychiatric symptoms. However, as of yet, the psychometric properties of
22 computational parameters have received limited attention (Ahn & Busemeyer, 2016; Browning

1 et al., 2020; Nair et al., 2020; Paulus et al., 2016). Model-agnostic measures derived from
2 reward processing tasks (e.g., percent correct) often have modest or poor reliability (Bland et
3 al., 2016; Enkavi et al., 2019; Plichta et al., 2012). The few studies that have examined the
4 reliability of PT and RL parameters based on gambling and various reward-processing tasks have
5 reported either poor-to-good reliability (Chung et al., 2017; Glockner & Pachur, 2012;
6 Scheibehenne & Pachur, 2015) or poor reliability (Moutoussis et al., 2018; Shahar et al., 2019).
7 However, recent studies have shown that higher reliability of RL processes can be achieved
8 using hierarchical estimation (Brown et al 2020; Waltmann et al., 2022). Computational
9 cognitive models are however often context-specific (Eckstein et al., 2022), and importantly, no
10 studies to date have reported the reliability of decision-making processes derived from a bandit
11 task with fluctuating (“restless”) reward and punishment probabilities (Daw et al., 2006) or an
12 individually calibrated gambling task (Charpentier et al., 2017), despite these showing promise
13 in computational psychiatry studies (e.g., Aylward et al., 2019; Charpentier et al., 2017).

14
15 A complementary perspective to understanding the reliability of computational cognitive
16 models can be obtained through prediction. Generative models offer a substantial advantage in
17 that they can both explain and predict behaviour. Unlike model-agnostic measures,
18 computational parameters fit to one dataset should be able to predict future behaviour in the
19 same individual. In other words, computational models can additionally be assessed by their
20 ability to forecast future behaviour, equivalent to out-of-sample validation. This type of
21 validation assesses model generalizability and is often referred to as predictive accuracy
22 (Busemeyer & Wang, 2000; Glockner & Pachur, 2012; Scheibehenne & Pachur, 2015), but it has

1 rarely been used as a metric of reliability. The aim of the current study was to assess the
2 reliability of model-agnostic and computational parameters derived from two widely-used
3 decision-making tasks (a restless four-armed bandit and a calibrated gambling task) using
4 standard measures of stability and reliability (respectively, practice effects and intraclass
5 correlations; ICCs) and additionally out-of-sample predictive accuracy for model parameters.

6

1 Methods and Materials

2 Participants

3 Fifty-four healthy participants were recruited from the UCL Institute of Cognitive Neuroscience

4 Subject Database. Four participants were excluded for failing to complete the second session

5 (final N=50: 32 females [64%]; age range=19-38; mean age=25.16, SD \pm 5.48 years; mean

6 education=17.38, SD \pm 3.24 years). Participants reported no current or past psychiatric or

7 neurological disorder, cannabis use in the past 31 days, alcohol consumption in the past 24

8 hours, or any other recreational drug use in the week prior to participation. Participants

9 provided written informed consent and were compensated at the end of their second session

10 with a flat rate of £30 and a bonus of up to £20 based on task winnings. The study was

11 approved by the UCL Psychology and Language Sciences Research Ethics Committee (Project ID

12 Number: fMRI/2013/005) and was performed in accordance with the Declaration of Helsinki.

13

14 Sample size was determined by an *a priori* power analysis in G*Power (Faul et al., 2007). The

15 power analysis was based on the smallest effect size of interest, $r=0.4$, since reliability below

16 this threshold is conventionally considered poor (Fleiss, 2011). Detecting an effect size of this

17 magnitude, at the one-tailed 0.05 alpha level with 90% power, requires 47 participants.

18

19 Study procedure and tasks

20 Participants completed a battery of computerized tasks, including a restless four-armed bandit

21 (Daw et al., 2006; Seymour et al., 2012) and an individually calibrated gambling task

1 (Charpentier et al., 2017) over two sessions (mean test-retest interval = 13.96 days, SD=0.20).

2 On each trial in the bandit task, participants chose one out of four bandits and received one out

3 of four possible outcomes: reward, punishment, neither reward nor punishment or both

4 reward and punishment (200 trials total). Win and loss probabilities fluctuated independently

5 over time and between boxes (Figure 1a-b; Supplemental Materials). In the gambling task,

6 participants chose between a 50-50 gamble and a sure option without receiving feedback. Trials

7 were classified as either mixed (50% chance to win or lose money gamble, or a sure option of 0

8 points) or gain-only (50% chance to win or receive nothing gamble, or a variable sure gain;

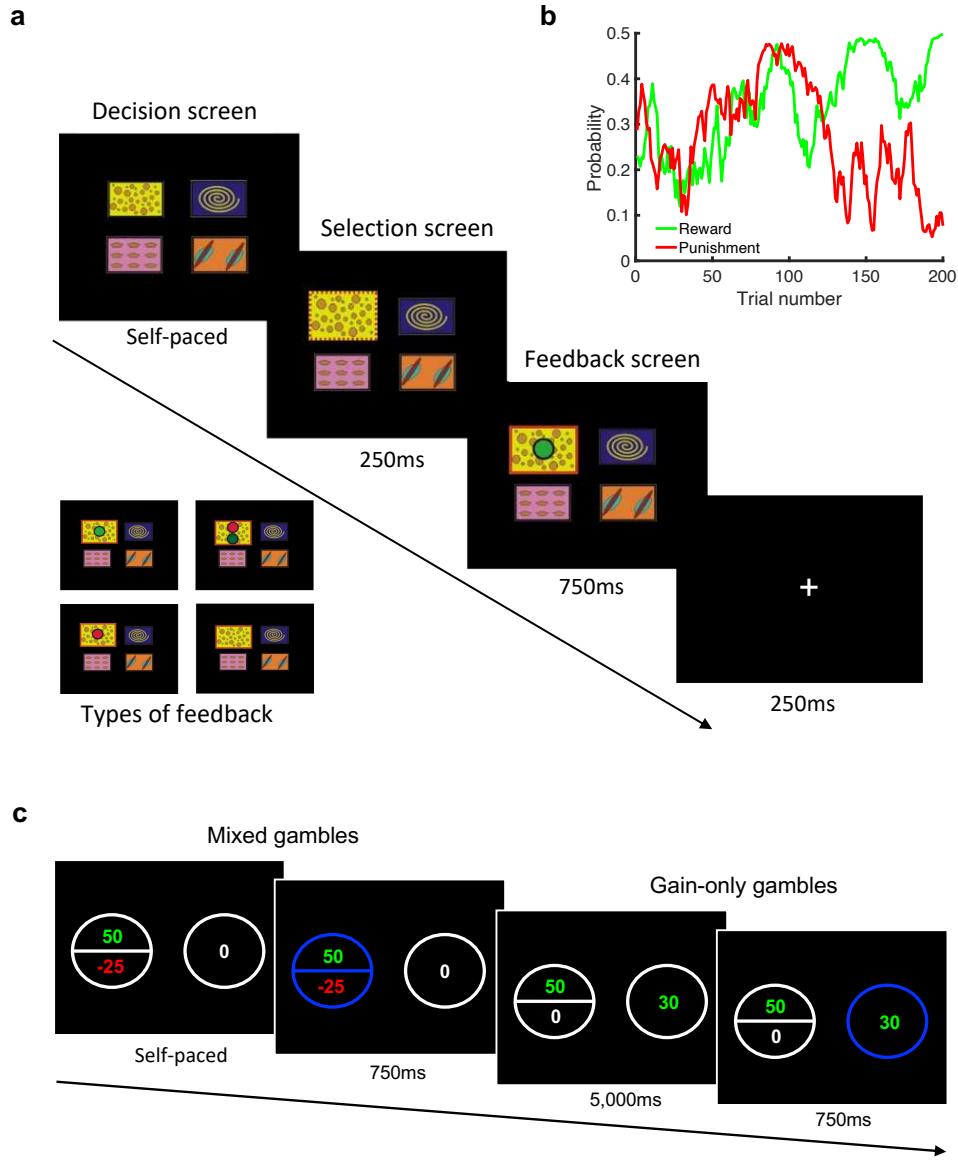
9 Figure 1c; Supplemental Materials). An initial training phase (50 mixed and 40 gain-only trials)

10 was used to create individually calibrated offers (centred on indifference points) in a second

11 phase (64 mixed and 56 gain-only trials). Calibration failed for one participant resulting in N=49

12 participants for this task. Both tasks lasted around 15 minutes.

13



1 Data analysis
2 Data were processed in Matlab (R2019b) and analysed in SPSS (v25, IBM Corp, Armonk, NY) and
3 R (v. 4.2.1). Computational modelling was performed with the hBayesDM package for R (v.
4 3.6.0; <https://github.com/CCS-Lab/hBayesDM>) (Ahn et al., 2017), which uses hierarchical
5 Bayesian modelling in Stan (v.2.21.2). For all analyses, $p < 0.05$ (two-tailed) was considered
6 statistically significant. Cohen's d_z (within-subject) effect sizes are reported for practice effects
7 (Lakens, 2013).

8

9 Model-agnostic task analyses
10 Model-agnostic measures for the bandit task included the mean probability of repeating a
11 choice after win-only, loss-only and no outcomes ('neither'). A repeated-measures analysis of
12 variance (ANOVA) was conducted with the within-subjects factors outcome (win, loss, neither)
13 and session (session 1, session 2) to assess basic behaviour and practice effects.

14

15 Model-agnostic measures for the gambling task included the mean probability of gambling on
16 mixed and gain-only trials. It was predicted that gambling would be higher on mixed trials,
17 which was assessed using a repeated-measures ANOVA with within-subjects factors gamble
18 (mixed, gain-only) and session (session 1, session 2).

19

20 As a sensitivity analysis, bandit and gamble model-agnostic measures were additionally derived
21 from trial-by-trial mixed-effects logistic regressions for each session since previous studies have

1 suggested that hierarchically predicted values improve reliability (Brown et al 2020; Waltmann
2 et al., 2022). Further details on all model-agnostic measures are in Supplemental Materials.

3

4 Computational modelling

5 The bandit task data was fit with seven different RL models previously described in detail
6 (Aylward et al., 2019) Supplemental Materials). Three PT models were fit to the gambling task
7 (Ahn et al., 2017; Charpentier et al., 2017; Kahneman & Tversky, 1979; Sokol-Hessner et al.,
8 2009). Modelling was conducted on the second phase of the gambling task (i.e., on individually
9 calibrated trials). The models were fit for each session separately, using separate hierarchical
10 priors (group-level parameters), as this has shown to provide more accurate fits (Valton et al.,
11 2020), and we wished to avoid artificially inflating reliability estimates. We also estimated
12 model fits under a single hierarchical prior (session 1 and session 2 data together) as a
13 sensitivity analysis (Supplemental Materials). Model comparison was performed with leave-
14 one-out information criterion (LOOIC) where the winning model was the one with the lowest
15 LOOIC. Several model validation checks were completed for the winning models, including
16 examination of MCMC convergences, parameter recovery and recapitulation of real data (Daw,
17 2011; Kruschke, 2015; Wilson & Collins, 2019; Supplemental Materials; Figure S1-S8).

18

19 Reliability analysis

20 Test-retest reliability was assessed with ICCs (ratios of intra-individual to inter-individual
21 variability (Koo & Li, 2016; McGraw & Wong, 1996)), with values of <0.40 interpreted as poor,
22 0.4-0.6 as fair, 0.6-0.75 as good, and >0.75 as excellent reliability (Fleiss, 2011). A two-way

1 mixed-effects model based on single-measures and absolute-agreement ICC was used (fixed
2 effect: testing time-interval, random effect: subject), equivalent to ICC(A,1) (McGraw & Wong,
3 1996).

4

5 We also estimated reliability directly within models as this has recently shown to increase
6 estimates by accounting for the uncertainty of measures (Brown et al 2020; Haines et al., 2020;
7 Waltmann et al., 2022). For model-agnostic measures, this was achieved by estimating a single
8 mixed-effects logistic regression with random intercepts and slopes, accounting data for both
9 sessions jointly by including session as a second-level grouping factor with subjects nested
10 within session (Brown et al 2020; Waltmann et al., 2022). This allows extracting variance
11 components from the logistic regression to calculate a one-way random-effects, absolute-
12 agreement, single-measure ICC: ICC(1) in the McGraw and Wong (1996) convention. This was
13 calculated for all bandit and gamble model-agnostic measures. For computational parameters,
14 model-calculated Pearson's r correlations were estimated between parameters from session 1
15 and 2 by fitting all the data from both sessions together and embedding a correlation matrix
16 between sessions in the winning RL and PT models (Haines et al., 2020; Pike et al., 2022). This is
17 different to the single-prior model, which does not include an embedded correlation matrix.

18

19 Posterior predictive performance
20 To assess to what extent an individual's future behaviour can be predicted using a generative
21 model fit to their own task performance two weeks earlier, we calculated the probability of
22 participants' choices on each trial (i.e., the softmax output), given their session 2 data and

1 model parameter estimates from session 1. Probabilities were averaged across trials for each
2 individual.

3

4 Since hierarchical parameter estimation produces ‘shrinkage’, effectively pulling parameter
5 estimates from different individuals closer to each other (which improves estimation accuracy),
6 it is possible that future performance may also be predicted above-chance using other
7 participants’ parameter estimates from session 1 (e.g., participant A’s parameter estimates
8 from session 1 predicting participant B’s session 2 choices). We therefore assessed whether
9 using an individual’s model parameter estimates from session 1 predicted the same individual’s
10 choices on session 2 better than using all other subjects’ model parameter estimates. To
11 construct the latter measure, for each subject, we predicted trial-by-trial choices on session 2
12 based on parameter estimates from every other participant’s session 1 model, and averaged
13 the probabilities across all participants. Additionally, we compared subjects’ own S1 parameters
14 in predicting S2 behaviour to the mean S1 model parameter priors in predicting future
15 behaviour.

16

17 Data accessibility

18 All script code and data are available on OSF at <https://osf.io/n7czx/>.

19

1 Results

2 Four-armed bandit task: model-agnostic results

3 *Basic behaviour and practice effects*

4 As expected, there was a main effect of outcome type on behaviour ($F_{(2,98)}=117.39, p<0.001$,
5 $\eta_p^2=0.71$; Figure 2a). The probability to repeat a choice was significantly greater after wins
6 compared with both losses and outcomes on which neither wins nor losses occurred, and
7 greater after neither compared with losses (all $p<0.001$). There was no significant main effect of
8 testing session ($F_{(1, 49)}=0.01, p=0.91, \eta_p^2<0.001$), but there was a significant outcome-by-session
9 interaction ($F_{(2,98)}=3.12, p=0.049, \eta_p^2=0.06$), reflecting slightly increased repeated choices after
10 wins and decreased repeated choices after losses. However, the difference in the tendency to
11 repeat a choice between session 1 and session 2 did not reach significance following any of the
12 outcome types (loss: $t_{(49)}=1.45, p=0.15, d_z=0.21$; win: $t_{(49)}=0.87, p=0.39, d_z=0.12$; neither:
13 $t_{(49)}=0.54, p=0.59, d_z=0.08$), and therefore we do not interpret this result further.

14

15 *Test-retest reliability*

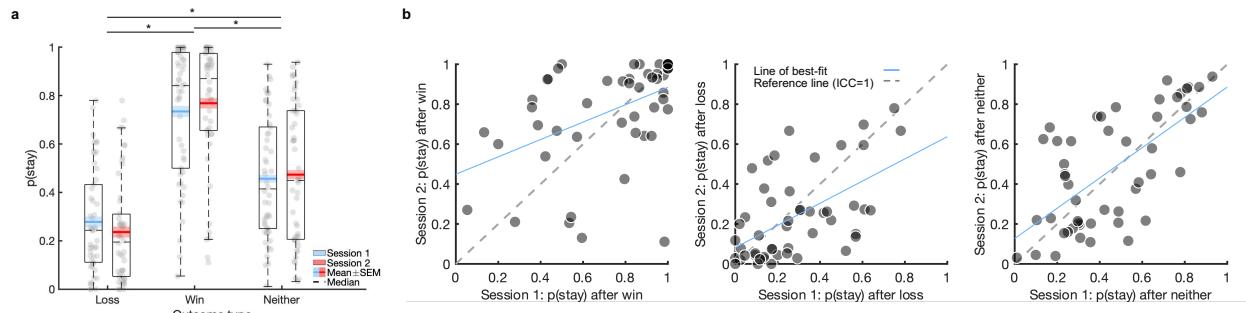
16 The model-agnostic measures exhibited fair-to-good reliability (Figure 2b; Table 1), which did
17 not improve substantially when examined in separate trial-by-trial mixed logistic regressions
18 (Table S1). However, reliability did increase, particularly for p(stay) after win and loss when
19 estimated as part of a joint mixed-effects logistic regression (Table 1).

20

21

22

1



2

3 **Figure 2: Basic behaviour, practice effects, and test-retest reliability of model-agnostic measures on the four-
4 armed bandit task.** Boxplots of the four-armed bandit task showing probability to stay after a certain outcome in
5 session 1 and 2 (a). The probability to stay was significantly different after each outcome type (Loss<Neither<Win)
6 but no clear practice effect was evident. Scatter plots of the model-agnostic measures comparing behaviour on
7 two testing sessions approximately 2 weeks apart (b). Lightly shaded regions in Figure 2a represent within-subjects
8 standard error of the mean (SEM). * $p<0.001$

9

10 Four-armed bandit task: modelling results
11 Model comparison indicated that the winning (most parsimonious) model was the five-
12 parameter 'Bandit4arm_lapse' (nomenclature from the hBayesDM package) model, with
13 reward and punishment learning rate parameters, reward and punishment sensitivity
14 parameters and a lapse parameter (final parameter captures random responding; Table S2),
15 consistent with previous reports (Aylward et al., 2019). Three individuals were excluded due to
16 difficulties in obtaining mean parameter estimates, as multiple peaks were evident in the
17 posterior distribution of at least one parameter (Figure S2). The Bandit4arm_lapse model was
18 therefore re-fit without these participants. Excluding these participants did not affect test-
19 retest reliability inference.

1

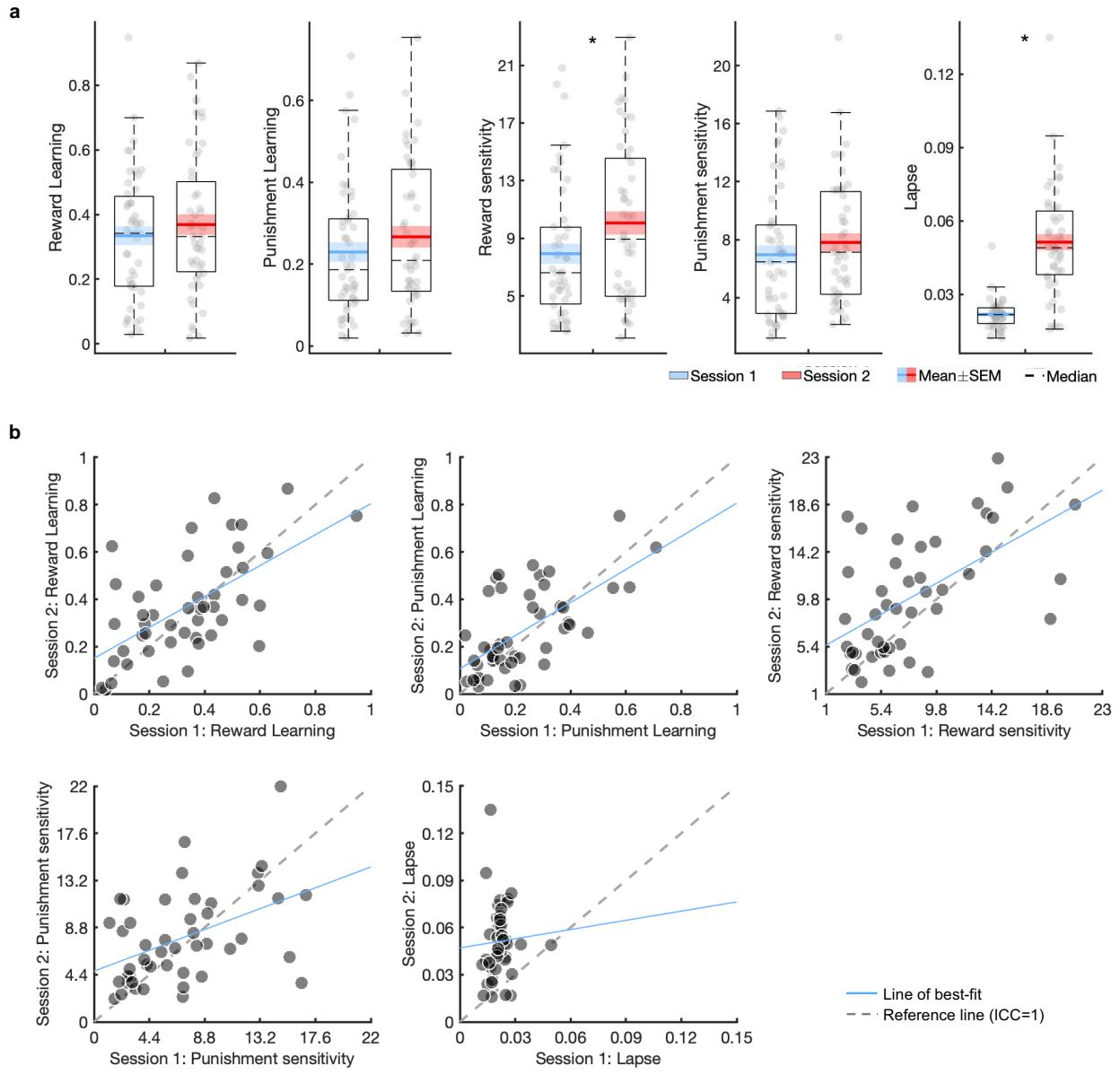
2 *Practice effects*

3 On session 2 there were significant increases in the reward sensitivity ($t_{(46)}=3.00, p=0.004$,
4 $d_z=0.44$) and lapse parameters ($t_{(46)}=8.88, p<0.001, d_z=1.29$), but not on any of the other
5 parameters (reward learning rate: $t_{(46)}=1.28, p=0.21, d_z=0.19$; punishment learning rate:
6 $t_{(46)}=1.74, p=0.09, d_z=0.25$; punishment sensitivity: $t_{(46)}=1.28, p=0.21, d_z=0.19$; Figure 3a).
7 However, there were no significant practice effects when the data was fit under a single
8 hierarchical prior (Table S3).

9

10 *Test-retest reliability*

11 All estimated Bandit4arm_lapse model parameters, except the lapse parameter, demonstrated
12 fair-to-good reliability (Figure 3b; Table 1), which did not substantially change when parameters
13 were estimated under a single hierarchical prior (Table S3). However, examining the correlation
14 between parameters as estimated within a generative model showed good-to-excellent
15 reliability, improving reliability across all but the lapse parameter (Table 1).



1 **Figure 3: Practice effects and test-retest reliability of the winning reinforcement learning model parameters**

2 derived from the four-armed bandit task. Boxplots show point estimates of the Bandit4arm_lapse model

3 parameters in session 1 and 2, fit under separate priors (a). Scatter plots of the Bandit4arm_lapse model

4 parameters over session 1 and 2 are presented (b). SEM: standard error of the mean. $*p < 0.05$.

5

6

Model-agnostic p(stay) measures (N=50)	ICC(A,1)	ICC(1)	Pearson's r
Summary statistics (Figure 2)			
Win	0.46 (0.21-0.65)	0.46 (0.21-0.65)	0.46 (0.20-0.65)
Loss	0.54 (0.32-0.71)	0.54 (0.31-0.71)	0.55 (0.32-0.72)
Neither	0.66 (0.48-0.79)	0.67 (0.48-0.80)	0.67 (0.48-0.80)
Hierarchical logistic regression (embedded reliability)			
Win		0.63	
Loss		0.63	
Neither		0.71	
Reinforcement learning model (N=47)	ICC(A,1)	ICC(1)	Pearson's r
Model estimated separately per session (Figure 3)			
Reward learning rate	0.60 (0.38-0.75)	0.60 (0.38-0.75)	0.60 (0.38-0.76)
Punishment learning rate	0.63 (0.42-0.77)	0.62 (0.41-0.77)	0.64 (0.43-0.78)
Reward sensitivity	0.52 (0.26-0.70)	0.50 (0.25-0.69)	0.56 (0.33-0.73)
Punishment sensitivity	0.45 (0.20-0.65)	0.45 (0.19-0.65)	0.46 (0.19-0.66)
Lapse	0.01 (-0.08-0.14)	-0.43 (-0.64- -0.17)	0.05 (-0.24-0.33)
Model estimated across sessions with embedded reliability			
Reward learning rate			0.71 (0.53-0.84)
Punishment learning rate			0.85 (0.69-0.95)
Reward sensitivity			0.68 (0.48-0.84)
Punishment sensitivity			0.64 (0.37-0.85)
Lapse			-0.01 (-0.65-0.68)

Table 1: Reliability of model-agnostic and computational measures of the four-armed bandit task.

All measures but the lapse parameter are significant at $p<0.05$. Brackets represent the 95% confidence interval.

- 1
- 2 *Posterior predictive performance*
- 3 Parameter estimates from session 1 predicted task performance on session 2 substantially
- 4 better than chance (mean = 42%, chance = 25% accuracy; $t_{(46)}=9.10$, $p<0.001$; Figure 4a),
- 5 indicating that the model could predict future choices by using a generative model fit to the
- 6 same participants' data two weeks earlier. Using an individual's parameter estimates to predict
- 7 their own future choices was significantly better than when that prediction was based on the
- 8 average of the other participants' session 1 estimates ($t_{(46)}=3.20$, $p=0.003$; Figure 4b). However,
- 9 there was no significant difference between using an individual's own session 1 parameter
- 10 estimates compared with the session 1 mean prior parameter in predicting future behaviour
- 11 ($t_{(46)}=1.04$, $p=0.30$; Figure 4c).

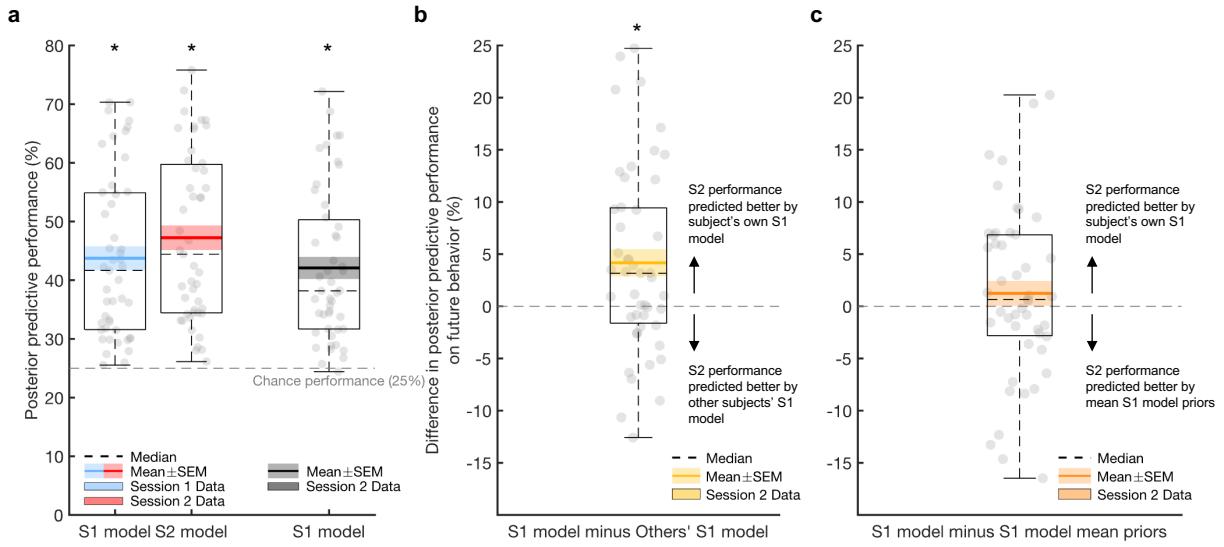


Figure 4: Posterior predictive performance of the winning reinforcement learning model derived from the four-armed bandit task. Boxplots depicting accuracy of bandit4arm_lapse model in predicting choices (a). Model estimates from session 1 (S1) predicted future session 2 (S2) behaviour above chance (black boxplot). Both S1 and S2 model estimates also predicted behaviour on the same session significantly above chance (blue and red boxplots). Predicting future performance (session 2 data) using a participant's own model parameter estimates was significantly better than using other participants' S1 model parameter estimates (b) but not when comparing against the mean S1 model priors (c). SEM: standard error of the mean. * $p<0.01$.

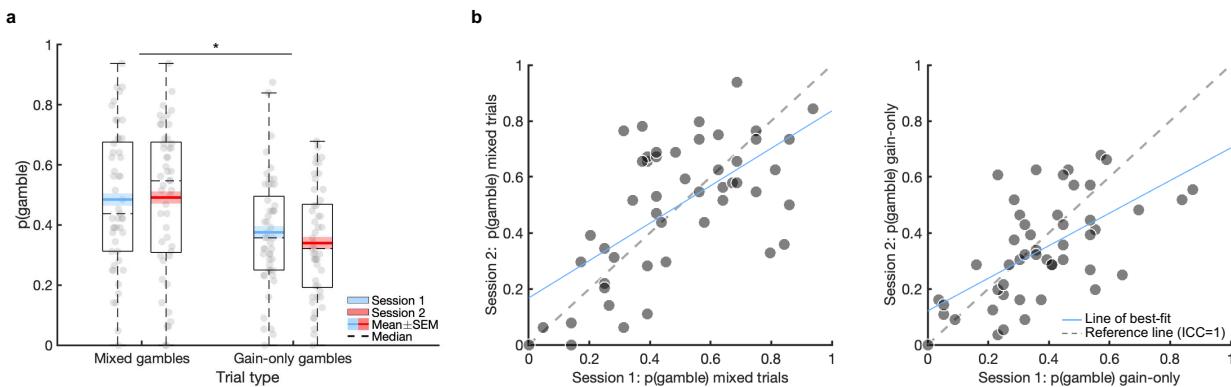
Gambling task: model-agnostic results
Basic behaviour and practice effects
 As expected, propensity to gamble was significantly higher on mixed gambles ($F_{(1, 48)}=13.71$, $p=0.001$, $\eta_p^2=0.22$). There were no significant main ($F_{(1, 48)}=0.76$, $p=0.40$, $\eta_p^2=0.02$) or interaction ($F_{(1, 48)}=1.07$, $p=0.31$, $\eta_p^2=0.02$) effects of session on the propensity to gamble (session

1 differences: probability to gamble on mixed trials $t_{(48)}=0.23$, $p=0.82$, $d_z=0.03$; probability to
2 gamble on gain-only trials $t_{(48)}=1.51$, $p=0.14$, $d_z=0.22$; Figure 5a).

3

4 *Test-retest reliability*

5 Model-agnostic measures on the gambling task exhibited fair-to-good reliability (Figure 5b;
6 Table 2) and did not change substantially when examined hierarchically in separate mixed
7 logistic regression (Table S4). Calculating reliability within a joint mixed logistic regression
8 numerically improved the reliability of the gambling model-agnostic measures (Table 2).



9
10 **Figure 5: Basic behaviour, practice effects, and test-retest reliability of model-agnostic measures on the**
11 **gambling task.** Boxplots show the probability to gamble based on the trial type in session 1 and 2, with no
12 significant session effects (a). Scatter plots of the model-agnostic measures over session 1 and 2 (b). Lightly shaded
13 regions in Figure 5a represent within-subjects standard error of the mean (SEM). * $p<0.001$.

14

15 Gambling task: modelling results

16 The winning model was the PT model ('ra_prospect' in the hBayesDM package) with loss
17 aversion, risk aversion and inverse temperature parameters (last parameter represents choice
18 consistency; Table S5), consistent with previous reports (Charpentier et al., 2017). A loss

1 aversion parameter above 1 represents overweighting of losses to gains, while a risk aversion
2 parameter less than 1 indicates aversion to risk. Neither test-retest nor practice effects were
3 substantially altered when the model was fit under a single hierarchical prior (Table S6).

4

5 *Practice effects*

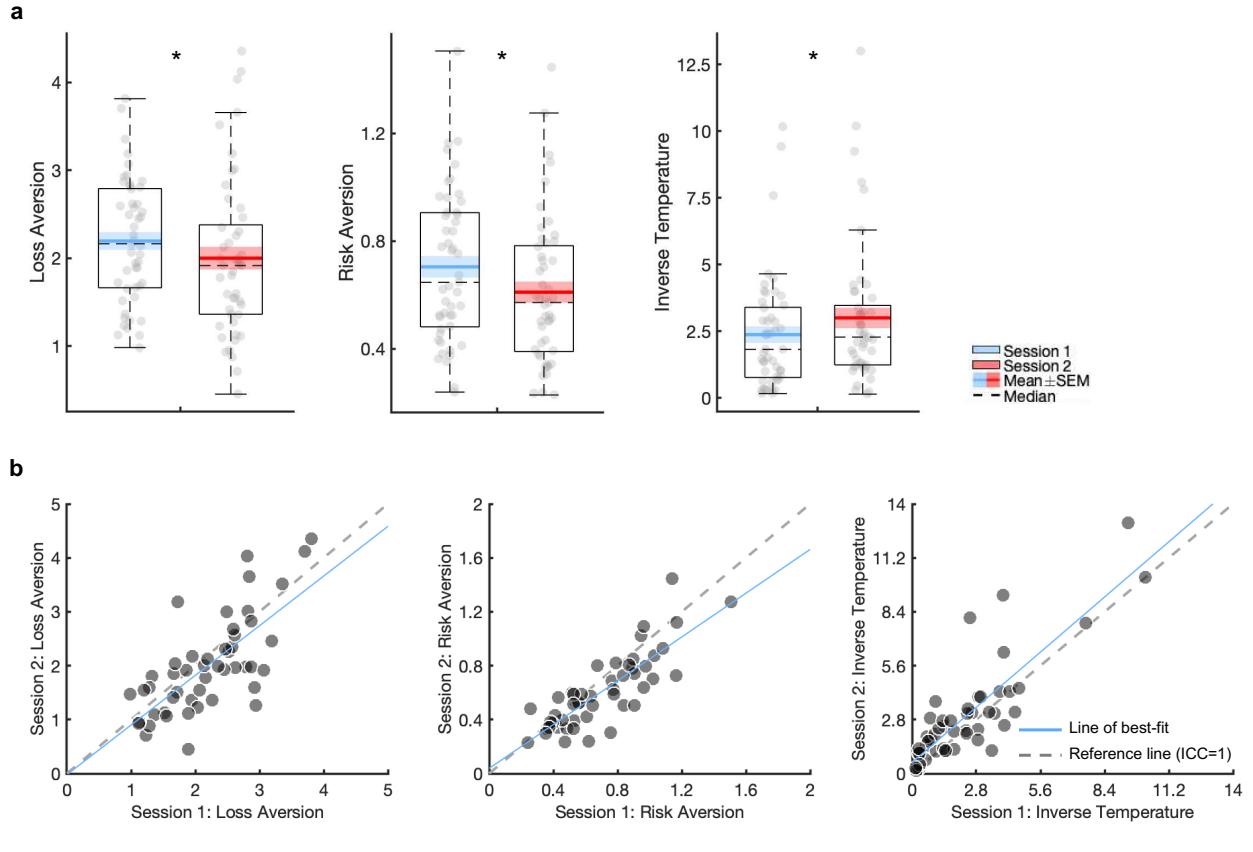
6 There were significant session effects on all PT model parameters (on session 2: decreased loss
7 aversion: $t_{(48)}=2.17, p=0.04, d_z=0.31$; decreased risk aversion: $t_{(48)}=4.04, p<0.001, d_z=0.58$;
8 increased inverse temperature: $t_{(48)}=3.07, p=0.004, d_z=0.44$; Figure 6a).

9

10 *Test-retest reliability*

11 All estimated parameters demonstrated good-to-excellent reliability (Figure 6b; Table 2), and
12 showed excellent reliability when estimating a correlation matrix embedded within the model
13 (Table 2).

14



1 **Figure 6: Practice effects and test-retest reliability of the prospect theory model derived from the gambling task.**

2 Boxplots show point estimates of the prospect theory model parameters in session 1 and 2, fit under separate
 3 priors (a). Scatter plots of the prospect theory model parameters over session 1 and 2 are presented (b). SEM:
 4 standard error of the mean. $*p < 0.05$.

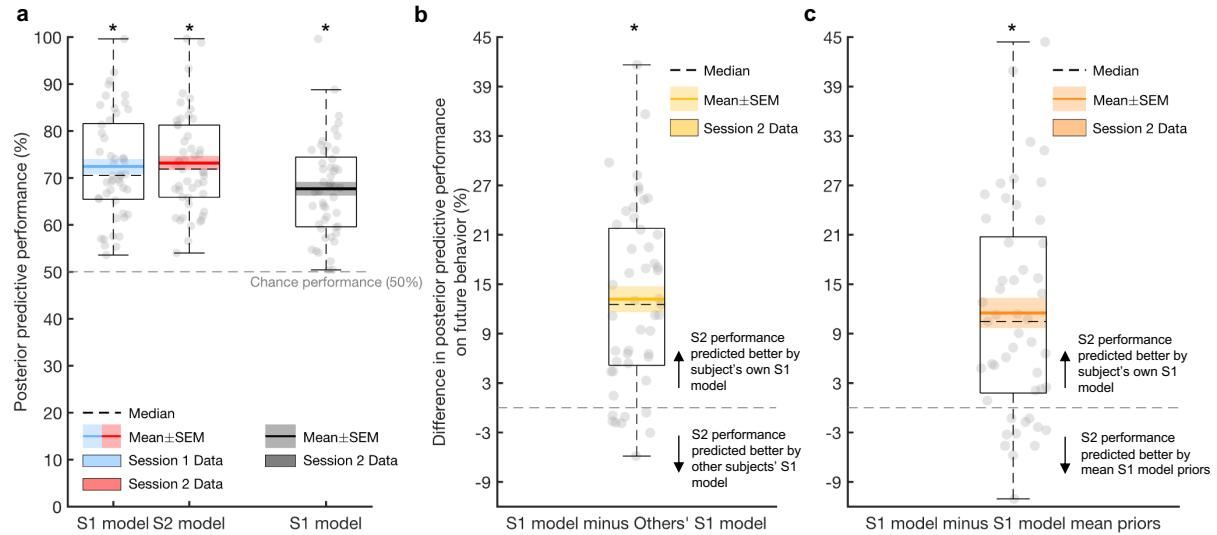
5

Model-agnostic p(gamble) measures	ICC(A,1)	ICC(1)	Pearson's r
Summary statistics (Figure 5)			
Mixed trials	0.63 (0.43-0.78)	0.63 (0.43-0.78)	0.63 (0.42-0.77)
Gain-only trials	0.59 (0.38-0.75)	0.59 (0.38-0.75)	0.60 (0.39-0.76)
Hierarchical logistic regression (embedded reliability)			
Mixed trials		0.73	
Gain-only trials		0.72	
Reinforcement learning model	ICC(A,1)	ICC(1)	Pearson's r
Model estimated separately per session (Figure 6)			
Loss aversion	0.68 (0.50-0.81)	0.68 (0.50-0.81)	0.72 (0.55-0.83)
Risk aversion	0.78 (0.55-0.89)	0.78 (0.64-0.87)	0.83 (0.71-0.90)
Inverse Temperature	0.80 (0.64-0.89)	0.80 (0.67-0.88)	0.84 (0.74-0.91)
Model estimated across sessions with embedded reliability			
Loss aversion			0.87 (0.77-0.94)
Risk aversion			0.90 (0.83-0.95)
Inverse Temperature			0.91 (0.85-0.96)

Table 2: Reliability of model-agnostic and computational measures of the gambling task.

All measures are significant at $p<0.05$. Brackets represent the 95% confidence interval.

- 1
- 2 *Posterior predictive performance*
- 3 PT model parameters from session 1 predicted future choices at session 2 significantly above
- 4 chance (mean = 68%, chance = 50% accuracy; $t_{(48)}= 12.08, p<0.001$; Figure 7a). Predicting future
- 5 performance at session 2 was significantly higher when based on participants' own parameter
- 6 estimates from session 1 compared with model parameter estimates of other participants from
- 7 session 1 ($t_{(48)}=8.38, p<0.001$; Figure 7b). This was also true when comparing between
- 8 participants' own session 1 parameter estimates with the mean session 1 prior for parameters
- 9 ($t_{(48)}=6.28, p<0.001$; Figure 7c).



1

2 **Figure 7: Posterior predictive performance of the prospect theory model derived from the gambling task.**

3 Boxplots depicting accuracy of prospect theory model in predicting choices (a). Session 1 (S1) model estimates
 4 predicted S1 behaviour significantly above chance (blue boxplot), as did session 2 (S2) model estimates on S2 data
 5 (red boxplot). Importantly, model parameter estimates from S1 predicted task performance from S2 above chance
 6 (black boxplot). Predicting future S2 performance using a participant's own S1 model parameter estimates (b) and mean S1 model priors (c).
 7 SEM: standard error of the mean. * $p<0.001$.
 8

9

10

1 Discussion

2 Reliability has garnered increased attention in recent years, with worryingly low reliability

3 across conventional measures from cognitive tasks and functional neuroimaging (Elliott et al.,

4 2020; Enkavi et al., 2019; Noble et al., 2019; Nord et al., 2017; Rodebaugh et al., 2016).

5 However, few attempts have been made to examine the reliability of computational cognitive

6 measures. Here we assessed the psychometric properties of computational models derived

7 from a restless four-armed bandit and a calibrated gambling task. Overall, most parameters

8 reflecting RL and decision-making processes exhibited adequate reliability and predicted future

9 performance well. These results provide promise for their use in clinical settings. However, this

10 conclusion depends on the specific parameters assessed in each task, highlighting the

11 complexities of translating tasks to the clinic.

12

13 Four-armed bandit RL model reliability

14 Reward and punishment learning rates from the bandit task demonstrated good reliability

15 while reward and punishment sensitivity showed fair reliability, suggesting that this task may be

16 more suitable for assessing learning rates than sensitivity. Elevated punishment learning rates

17 (faster learning in the face of negative outcomes) and lapse values have been associated with

18 greater mood and anxiety symptoms, representing potential measurable mechanistic treatment

19 targets (Aylward et al., 2019). However, the present study suggests that the lapse parameter,

20 which exhibited poor reliability as assessed by the bandit task, may not be a suitable target.

21 This parameter measures responding not captured by the model (including goal-directed and

22 random exploration), and the sources of this 'noise' might differ across sessions. It is therefore

1 perhaps unsurprising that this parameter was unreliable. Crucially, the lapse parameter showed
2 poor recoverability, which places an upper limit on its potential reliability. Some of this poor
3 recoverability may be explained by limited lapse variation, especially in session 1 (Figure S3).
4 The distribution of the group-level standard deviation lapse parameter was biased towards
5 smaller values here such that the lapse parameter did not vary greatly between individuals
6 (Figure S5). This suggests that the lapse parameter could be replaced with a constant and
7 inference on this parameter is not advised.

8

9 Although no prior studies have specifically investigated ICC properties of the current RL model,
10 one previous study found similarly poor reliability of the lapse parameter across six months in a
11 go/no-go RL model in adolescents (Moutoussis et al., 2018). In contrast to our results, this study
12 also reported poor reliability of both reward and punishment learning rates. These differences
13 may arise for a multitude of reason, such as using different tasks (an orthogonalised go/no-go
14 task versus a restless bandit task), testing time-windows (six months versus two weeks),
15 populations (adolescents versus adults), or models. It is not possible to delineate these
16 diverging results without systematically comparing these factors in one study. Importantly,
17 however, we provide evidence that it is possible to achieve at least moderate reliability for
18 some canonical RL parameters.

19

20 Interestingly, the model-agnostic outcome measures of the bandit task exhibited similar
21 reliability to the computational measures. Model-agnostic measures of cognitive tasks have
22 often been reported to exhibit poor-to-moderate reliability (Enkavi et al., 2019; Hedge et al.,

1 2018; Plichta et al., 2012; Rodebaugh et al., 2016). It has been argued that this may be due to
2 their inability to capture the generative process underlying task performance (Huys et al., 2021;
3 Price et al., 2019). Our results suggest that it should not be assumed that computational
4 parameters will always provide greater reliability than non-computational ones. However, the
5 model-agnostic outcome measures are only a proxy of the processes the bandit task assesses,
6 as it is difficult to compute model-agnostic equivalents of some parameters, such as
7 reward/punishment sensitivity. Indeed, models make explicit and falsifiable predictions of the
8 components driving behaviour, which can be refined and used to simulate artificial data to
9 generate new predictions. Thus, computational modelling is a more rigorous and preferable
10 method for assessing behaviour than model-agnostic measures, which unlike computational
11 methods, lack the mechanistic insights into the underlying processes generating behaviour.

12

13 Gambling PT model reliability
14 While reliability of parameters ranged from poor-to-good in the bandit task, parameters from
15 the gambling task showed good-to-excellent test-retest reliability. These were also higher than
16 the reliability of the model-agnostic measures, suggesting that computational models may offer
17 advantages in psychometric properties here. In particular, the risk aversion parameter, which
18 has previously been associated with anxiety (Charpentier et al., 2017), exhibited excellent
19 reliability ($ICC=0.78$), providing promise for use in clinical research. These results show higher
20 reliability than previous studies (loss aversion $r\approx0.25-0.61$, risk aversion $r\approx0.50-0.60$, inverse
21 temperature $r\approx0.30-0.60$ [Chung et al., 2017; Glockner & Pachur, 2012; Scheibehenne & Pachur,
22 2015]). These studies all used different estimation procedures, including hierarchical Bayesian,

1 and employed both longer and shorter testing time-windows than the current study, suggesting
2 that these factors may not fully explain the differences. It is possible that our results instead
3 stem from different PT model specifications, as well as different task designs. Indeed, a strength
4 of the gambling task is that we calibrated offers to each individual's indifference point
5 (Charpentier et al., 2017). A similar approach of dynamically updating parameter values to each
6 individual during task performance has previously been suggested as a solution to unreliable
7 cognitive tasks (Palminteri & Chevallier, 2018). This may allow for removing any potential state
8 influences across participants to extract more trait-like measures of the variables of interest
9 (here risk/loss aversion).

10

11 Predictive accuracy
12 We also examined how well the models predicted future task performance, which provides
13 complementary perspective on reliability, unique to computationally-informed measures.
14 Notably, for the PT model participants' own parameter estimates from the first session were
15 better at predicting their future performance compared with using parameter estimates from
16 all other participants and from model priors. Individuals' own RL parameters only provided an
17 advantage in predicting future performance when compared with other's parameter estimates
18 but not model priors. It is likely that the RL model did not perform as well on this metric as the
19 PT model because the RL model does not provide as close a fit to behaviour, potentially due to
20 some participants performing at chance level. Overall, this indicates that individuals may also
21 reliably differ in the cognitive mechanisms underlying their decisions, and offers reassurance
22 that hierarchical estimation procedures are suitable for inter-individual inferences (Brown et

1 al., 2020; Daw, 2011; Scheibehenne & Pachur, 2015). In other words, individuals show relatively
2 unique computational decision-making profiles, particularly in the PT model. This is consistent
3 with two previous studies using a different PT model and gambling task (Glockner & Pachur,
4 2012; Scheibehenne & Pachur, 2015).

5

6 Implications

7 The RL parameters showed relatively modest reliability, suggesting that these processes are
8 more vulnerable to state influences than PT parameters. A consequence of this is that larger
9 sample sizes may be required to examine effects, as effect sizes would be expected to be lower,
10 relative to PT tasks (Fleiss, 2011). Interpreting the marked difference in reliability between the
11 PT and RL models is not straightforward, as these models measure complementary aspects of
12 cognition. The bandit task is a learning paradigm that requires constant updating of optimal
13 choices. It is possible that in the first session individuals had not yet stabilized on a cognitive
14 strategy and were still learning the task structure, as indicated by lower evidence of the winning
15 model in session 1 compared with session 2 (Table S2). It would be interesting to explore if an
16 initial baseline session, would improve test-retest reliability. It should be noted that we found
17 substantially greater reliability of RL (good-to-excellent) and PT (excellent) parameters when
18 estimated within the generative model. These provide an upper theoretical reliability bound of
19 a parameter and are mainly relevant for within-subjects applications (if changes are modelled
20 within a model). In addition, although previous studies suggest that hierarchically measured
21 model-agnostic predicted values substantially improve reliability (Brown et al., 2020; Waltmann
22 et al., 2020), this was not the case in the current study. However, calculating the reliability of

1 model-agnostic measures directly from a joint logistic regression improved model-agnostic
2 reliabilities but measures remained in the range of good reliability.

3

4 Limitations

5 A potential limitation of our study is the sample tested, as the reliability of tasks in healthy
6 individuals may differ from that in clinical groups. Similarly, our results only speak to reliability
7 over two weeks. Thus, it is possible that longer time periods may produce lower reliability,
8 which should be assessed in future studies. Reliability over two weeks is particularly informative
9 for interventional studies such as randomised controlled trials (e.g., for rapid-acting
10 antidepressants or for early markers of response for traditional antidepressants/
11 psychotherapies). This time-window is also in line with other reliability studies aiming to
12 establish reliability of measures for e.g., individual differences (e.g., Hedge et al., 2018; Nord et
13 al., 2017), based on the assumption that measures should remain relatively stable over a short
14 time-period.

15

16 Conclusion

17 In summary, we show that commonly-used computational parameters derived from an RL
18 ‘restless’ bandit task and a calibrated gambling task exhibit fair-to-excellent reliability.
19 Specifically, learning rates showed good reliability and sensitivity parameters showed fair
20 reliability from the RL model, while loss aversion had good reliability and risk aversion and
21 inverse temperature displayed excellent reliability from the PT model. These models can
22 further be used to predict future behaviour in the same individuals, especially PT model

- 1 parameters, indicating that the decision-making processes assessed in these tasks represent
- 2 relatively consistent and unique characteristics of an individual. These findings take us one step
- 3 closer to translating computational measures of behaviour into clinical application.

1 Acknowledgements

2 This work was supported by a Wellcome Trust - NIH PhD studentship (200934/Z/16/Z) to A.M.;
3 a National Institute for Health Research (NIHR) Biomedical Research Center (BRC) fellowship to
4 V.V.; and by a Wellcome Investigator Award (101798/Z/13/Z) to J.P.R. The funders had no role
5 in the study design, data collection and analysis, decision to publish or preparation of the
6 manuscript. For the purpose of Open Access, the author has applied a CC BY public copyright
7 licence to any Author Accepted Manuscript version arising from this submission. The authors
8 thank Eoin Travers and Oliver J. Robinson for valuable input on an earlier version of the
9 manuscript. A preprint was previously published at bioRxiv
10 (<https://doi.org/10.1101/2021.06.30.450026>).

11

12 Competing interests

13 The author(s) has/have no competing interests to declare.

14

15 Authors' contribution

16 A.M. and J.P.R. conceived and designed the study. A.M. acquired the data and performed
17 analyses under the supervision of V.V. and J.P.R. All authors contributed to interpreting the
18 data and drafting or substantially revising of the manuscript.

1 References

2 Adams, R. A., Huys, Q. J., & Roiser, J. P. (2016). Computational Psychiatry: towards a
3 mathematically informed understanding of mental illness. *J Neurol Neurosurg
4 Psychiatry*, 87(1), 53-63. <https://doi.org/10.1136/jnnp-2015-310737>

5 Ahn, W. Y., & Busemeyer, J. R. (2016). Challenges and promises for translating computational
6 tools into clinical practice. *Curr Opin Behav Sci*, 11, 1-7.
7 <https://doi.org/10.1016/j.cobeha.2016.02.001>

8 Ahn, W. Y., Haines, N., & Zhang, L. (2017). Revealing Neurocomputational Mechanisms of
9 Reinforcement Learning and Decision-Making With the hBayesDM Package. *Comput
10 Psychiatr*, 1, 24-57. https://doi.org/10.1162/CPSY_a_00002

11 Aylward, J., Valton, V., Ahn, W. Y., Bond, R. L., Dayan, P., Roiser, J. P., & Robinson, O. J. (2019).
12 Altered learning under uncertainty in unmedicated mood and anxiety disorders. *Nat
13 Hum Behav*, 3, 1116–1123. <https://doi.org/10.1038/s41562-019-0628-0>

14 Baek, K., Kwon, J., Chae, J. H., Chung, Y. A., Kralik, J. D., Min, J. A., Huh, H., Choi, K. M., Jang, K. I.,
15 Lee, N. B., Kim, S., Peterson, B. S., & Jeong, J. (2017). Heightened aversion to risk and
16 loss in depressed patients with a suicide attempt history. *Sci Rep*, 7(1), 11228.
17 <https://doi.org/10.1038/s41598-017-10541-5>

18 Bland, A. R., Roiser, J. P., Mehta, M. A., Schei, T., Boland, H., Campbell-Meiklejohn, D. K.,
19 Emsley, R. A., Munafo, M. R., Penton-Voak, I. S., Seara-Cardoso, A., Viding, E., Voon, V.,
20 Sahakian, B. J., Robbins, T. W., & Elliott, R. (2016). EMOTICOM: A Neuropsychological
21 Test Battery to Evaluate Emotion, Motivation, Impulsivity, and Social Cognition. *Front
22 Behav Neurosci*, 10, 25. <https://doi.org/10.3389/fnbeh.2016.00025>

1 Brown, J. K., Waltz, J. A., Strauss, G. P., McMahon, R. P., Frank, M. J., & Gold, J. M. (2013).

2 Hypothetical decision making in schizophrenia: the role of expected value computation

3 and "irrational" biases. *Psychiatry Res*, 209(2), 142-149.

4 <https://doi.org/10.1016/j.psychres.2013.02.034>

5 Brown, V. M., Chen, J., Gillan, C. M., & Price, R. B. (2020). Improving the Reliability of

6 Computational Analyses: Model-Based Planning and Its Relationship With Compulsivity.

7 *Biol Psychiatry Cogn Neurosci Neuroimaging*, 5(6), 601-609.

8 <https://doi.org/10.1016/j.bpsc.2019.12.019>

9 Browning, M., Carter, C. S., Chatham, C., Den Ouden, H., Gillan, C. M., Baker, J. T., Chekroud, A.

10 M., Cools, R., Dayan, P., Gold, J., Goldstein, R. Z., Hartley, C. A., Kepecs, A., Lawson, R. P.,

11 Mourao-Miranda, J., Phillips, M. L., Pizzagalli, D. A., Powers, A., Rindskopf, D., . . . Paulus,

12 M. (2020). Realizing the Clinical Potential of Computational Psychiatry: Report From the

13 Banbury Center Meeting, February 2019. *Biol Psychiatry*, 88(2), e5-e10.

14 <https://doi.org/10.1016/j.biopsych.2019.12.026>

15 Busemeyer, J. R., & Wang, Y. M. (2000). Model Comparisons and Model Selections Based on

16 Generalization Criterion Methodology. *J Math Psychol*, 44(1), 171-189.

17 <https://doi.org/10.1006/jmps.1999.1282>

18 Charpentier, C. J., Aylward, J., Roiser, J. P., & Robinson, O. J. (2017). Enhanced Risk Aversion,

19 But Not Loss Aversion, in Unmedicated Pathological Anxiety. *Biol Psychiatry*, 81(12),

20 1014-1022. <https://doi.org/10.1016/j.biopsych.2016.12.010>

1 Charpentier, C. J., De Martino, B., Sim, A. L., Sharot, T., & Roiser, J. P. (2016). Emotion-induced
2 loss aversion and striatal-amygdala coupling in low-anxious individuals. *Soc Cogn Affect
3 Neurosci*, 11(4), 569-579. <https://doi.org/10.1093/scan/nsv139>

4 Chung, D., Kadlec, K., Aimone, J. A., McCurry, K., King-Casas, B., & Chiu, P. H. (2017). Valuation
5 in major depression is intact and stable in a non-learning environment. *Sci Rep*, 7,
6 44374. <https://doi.org/10.1038/srep44374>

7 Daw, N. D. (2011). Trial-by-trial data analysis using computational models. In *Decision Making,
8 Affect, and Learning: Attention and Performance XXIII*. Oxford University Press.
9 <https://doi.org/https://doi.org/10.1093/acprof:oso/9780199600434.003.0001>

10 Daw, N. D., & Doya, K. (2006). The computational neurobiology of learning and reward. *Curr
11 Opin Neurobiol*, 16(2), 199-204. <https://doi.org/10.1016/j.conb.2006.03.006>

12 Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for
13 exploratory decisions in humans. *Nature*, 441(7095), 876-879.
14 <https://doi.org/10.1038/nature04766>

15 Dayan, P., & Niv, Y. (2008). Reinforcement learning: the good, the bad and the ugly. *Curr Opin
16 Neurobiol*, 18(2), 185-196. <https://doi.org/10.1016/j.conb.2008.08.003>

17 *Diagnostic and statistical manual of mental disorders: DSM-5*. (2013). (Fifth ed.). American
18 Psychiatric Association.

19 Eckstein, M. K., Master, S. L., Xia, L., Dahl, R. E., Wilbrecht, L., & Collins, A. G. E. (2022). The
20 Interpretation of Computational Model Parameters Depends on the Context. Preprint at
21 bioRxiv <https://doi.org/10.1101/2021.05.28.446162>

1 Elliott, M. L., Knott, A. R., Ireland, D., Morris, M. L., Poulton, R., Ramrakha, S., Sison, M. L.,

2 Moffitt, T. E., Caspi, A., & Hariri, A. R. (2020). What Is the Test-Retest Reliability of

3 Common Task-Functional MRI Measures? New Empirical Evidence and a Meta-Analysis.

4 *Psychol Sci*, 31(7), 792-806. <https://doi.org/10.1177/0956797620916786>

5 Enkavi, A. Z., Eisenberg, I. W., Bissett, P. G., Mazza, G. L., MacKinnon, D. P., Marsch, L. A., &

6 Poldrack, R. A. (2019). Large-scale analysis of test-retest reliabilities of self-regulation

7 measures. *Proc Natl Acad Sci U S A*, 116(12), 5472-5477.

8 <https://doi.org/10.1073/pnas.1818430116>

9 Eshel, N., & Roiser, J. P. (2010). Reward and punishment processing in depression. *Biol*

10 *Psychiatry*, 68(2), 118-124. <https://doi.org/10.1016/j.biopsych.2010.01.027>

11 Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible statistical power

12 analysis program for the social, behavioral, and biomedical sciences. *Behav Res*

13 *Methods*, 39(2), 175-191. <https://doi.org/10.3758/bf03193146>

14 Fleiss, J. L. (2011). Reliability of Measurement. In *The Design and Analysis of Clinical*

15 *Experiments* (pp. 1-32). <https://doi.org/10.1002/9781118032923.ch1>

16 Glockner, A., & Pachur, T. (2012). Cognitive models of risky choice: parameter stability and

17 predictive accuracy of prospect theory. *Cognition*, 123(1), 21-32.

18 <https://doi.org/10.1016/j.cognition.2011.12.002>

19 Hadlaczky, G., Hokby, S., Mkrtchian, A., Wasserman, D., Balazs, J., Machin, N., Sarchiapone, M.,

20 Sisask, M., & Carli, V. (2018). Decision-Making in Suicidal Behavior: The Protective Role

21 of Loss Aversion. *Front Psychiatry*, 9, 116. <https://doi.org/10.3389/fpsyg.2018.00116>

1 Haines, N., Kvam, P., Irving, L., Smith, C., Beauchaine, T. P., Pitt, M. A., Ahn, W.-Y., & Turner, B.
2 (2020). Learning from the Reliability Paradox: How Theoretically Informed Generative
3 Models Can Advance the Social, Behavioral, and Brain Sciences. *PsyArXiv*.
4 <https://doi.org/10.31234/osf.io/xr7y3>

5 Hartley, C. A., & Phelps, E. A. (2012). Anxiety and decision-making. *Biol Psychiatry*, 72(2), 113-
6 118. <https://doi.org/10.1016/j.biopsych.2011.12.027>

7 Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks
8 do not produce reliable individual differences. *Behav Res Methods*, 50(3), 1166-1186.
9 <https://doi.org/10.3758/s13428-017-0935-1>

10 Huys, Q. J. M. (2018). Advancing Clinical Improvements for Patients Using the Theory-Driven
11 and Data-Driven Branches of Computational Psychiatry. *JAMA Psychiatry*, 75(3), 225-
12 226. <https://doi.org/10.1001/jamapsychiatry.2017.4246>

13 Huys, Q. J. M., Browning, M., Paulus, M. P., & Frank, M. J. (2021). Advances in the
14 computational understanding of mental illness. *Neuropsychopharmacology*, 46(1), 3-19.
15 <https://doi.org/10.1038/s41386-020-0746-4>

16 Huys, Q. J. M., Maia, T. V., & Frank, M. J. (2016). Computational psychiatry as a bridge from
17 neuroscience to clinical applications. *Nat Neurosci*, 19(3), 404-413.
18 <https://doi.org/10.1038/nn.4238>

19 Huys, Q. J. M., Moutoussis, M., & Williams, J. (2011). Are computational models of any use to
20 psychiatry? *Neural Netw*, 24(6), 544-551. <https://doi.org/10.1016/j.neunet.2011.03.001>

21 Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk.
22 *Econometrica*, 47(2), 263-291. <https://doi.org/10.2307/1914185>

1 Klaus, F., Chumbley, J. R., Seifritz, E., Kaiser, S., & Hartmann-Riemer, M. (2020). Loss Aversion
2 and Risk Aversion in Non-Clinical Negative Symptoms and Hypomania. *Front Psychiatry*,
3 11, 574131. <https://doi.org/10.3389/fpsyg.2020.574131>

4 Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation
5 Coefficients for Reliability Research. *J Chiropr Med*, 15(2), 155-163.
6 <https://doi.org/10.1016/j.jcm.2016.02.012>

7 Kruschke, J. K. (2015). *Doing Bayesian Data Analysis: A tutorial with R, JAGS, and Stan* (Second
8 ed.). Academic Press.

9 Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a
10 practical primer for t-tests and ANOVAs [Review]. *Frontiers in Psychology*, 4(863), 863.
11 <https://doi.org/10.3389/fpsyg.2013.00863>

12 Maia, T. V. (2009). Reinforcement learning, conditioning, and the brain: Successes and
13 challenges. *Cogn Affect Behav Neurosci*, 9(4), 343-364.
14 <https://doi.org/10.3758/CABN.9.4.343>

15 Maia, T. V., & Frank, M. J. (2011). From reinforcement learning models to psychiatric and
16 neurological disorders. *Nat Neurosci*, 14(2), 154-162. <https://doi.org/10.1038/nn.2723>

17 McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation
18 coefficients. *Psychological Methods*, 1(1), 30-46. <https://doi.org/10.1037/1082-989x.1.1.30>

20 Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends
21 Cogn Sci*, 16(1), 72-80. <https://doi.org/10.1016/j.tics.2011.11.018>

1 Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational roles for dopamine in
2 behavioural control. *Nature*, 431(7010), 760-767. <https://doi.org/10.1038/nature03015>

3 Moutoussis, M., Bullmore, E. T., Goodyer, I. M., Fonagy, P., Jones, P. B., Dolan, R. J., Dayan, P., &
4 Neuroscience in Psychiatry Network Research, C. (2018). Change, stability, and
5 instability in the Pavlovian guidance of behaviour from adolescence to young adulthood.
6 *PLoS Comput Biol*, 14(12), e1006679. <https://doi.org/10.1371/journal.pcbi.1006679>

7 Nair, A., Rutledge, R. B., & Mason, L. (2020). Under the Hood: Using Computational Psychiatry
8 to Make Psychological Therapies More Mechanism-Focused. *Front Psychiatry*, 11, 140.
9 <https://doi.org/10.3389/fpsyg.2020.00140>

10 Niv, N. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, 53(3),
11 139-154.

12 Noble, S., Scheinost, D., & Constable, R. T. (2019). A decade of test-retest reliability of
13 functional connectivity: A systematic review and meta-analysis. *Neuroimage*, 203,
14 116157. <https://doi.org/10.1016/j.neuroimage.2019.116157>

15 Nord, C. L., Gray, A., Charpentier, C. J., Robinson, O. J., & Roiser, J. P. (2017). Unreliability of
16 putative fMRI biomarkers during emotional face processing. *Neuroimage*, 156, 119-127.
17 <https://doi.org/10.1016/j.neuroimage.2017.05.024>

18 Palminteri, S., & Chevallier, C. (2018). Can We Infer Inter-Individual Differences in Risk-Taking
19 From Behavioral Tasks? *Front Psychol*, 9, 2307.
20 <https://doi.org/10.3389/fpsyg.2018.02307>

1 Palminteri, S., Wyart, V., & Koechlin, E. (2017). The Importance of Falsification in Computational
2 Cognitive Modeling. *Trends Cogn Sci*, 21(6), 425-433.
3 <https://doi.org/10.1016/j.tics.2017.03.011>

4 Patzelt, E. H., Hartley, C. A., & Gershman, S. J. (2018). Computational Phenotyping: Using
5 Models to Understand Individual Differences in Personality, Development, and Mental
6 Illness. *Personal Neurosci*, 1, e18. <https://doi.org/10.1017/pen.2018.14>

7 Paulus, M. P., Huys, Q. J., & Maia, T. V. (2016). A Roadmap for the Development of Applied
8 Computational Psychiatry. *Biol Psychiatry Cogn Neurosci Neuroimaging*, 1(5), 386-392.
9 <https://doi.org/10.1016/j.bpsc.2016.05.001>

10 Paulus, M. P., & Thompson, W. K. (2019). Computational approaches and machine learning for
11 individual-level treatment predictions. *Psychopharmacology (Berl)*.
12 <https://doi.org/10.1007/s00213-019-05282-4>

13 Pike, A. C., Tan, K., Ansari, H. J., Wing, M., & Robinson, O. J. (2022). Test-retest reliability of
14 affective bias tasks. Preprint at PsyArXiv. <https://doi.org/10.31234/osf.io/n2fkh>

15 Plichta, M. M., Schwarz, A. J., Grimm, O., Morgen, K., Mier, D., Haddad, L., Gerdes, A. B., Sauer,
16 C., Tost, H., Esslinger, C., Colman, P., Wilson, F., Kirsch, P., & Meyer-Lindenberg, A.
17 (2012). Test-retest reliability of evoked BOLD signals from a cognitive-emotive fMRI test
18 battery. *Neuroimage*, 60(3), 1746-1758.
19 <https://doi.org/10.1016/j.neuroimage.2012.01.129>

20 Price, R. B., Brown, V., & Siegle, G. J. (2019). Computational Modeling Applied to the Dot-Probe
21 Task Yields Improved Reliability and Mechanistic Insights. *Biol Psychiatry*, 85(7), 606-
22 612. <https://doi.org/10.1016/j.biopsych.2018.09.022>

1 Rodebaugh, T. L., Scullin, R. B., Langer, J. K., Dixon, D. J., Huppert, J. D., Bernstein, A., Zvielli, A.,
2 & Lenze, E. J. (2016). Unreliability as a threat to understanding psychopathology: The
3 cautionary tale of attentional bias. *J Abnorm Psychol*, 125(6), 840-851.
4 <https://doi.org/10.1037/abn0000184>

5 Ruggeri, K., Ali, S., Berge, M. L., Bertoldo, G., BJORNDAL, L. D., Cortijos-Bernabeu, A., Davison, C.,
6 Demic, E., Esteban-Serna, C., Friedemann, M., Gibson, S. P., Jarke, H., Karakasheva, R.,
7 Khorrami, P. R., Kveder, J., Andersen, T. L., Lofthus, I. S., McGill, L., Nieto, A. E., . . . Folke,
8 T. (2020). Replicating patterns of prospect theory for decision under risk. *Nat Hum
Behav*, 4(6), 622-633. <https://doi.org/10.1038/s41562-020-0886-x>

10 Scheibehenne, B., & Pachur, T. (2015). Using Bayesian hierarchical parameter estimation to
11 assess the generalizability of cognitive models of choice. *Psychon Bull Rev*, 22(2), 391-
12 407. <https://doi.org/10.3758/s13423-014-0684-4>

13 Schonberg, T., Fox, C. R., & Poldrack, R. A. (2011). Mind the gap: bridging economic and
14 naturalistic risk-taking with cognitive neuroscience. *Trends Cogn Sci*, 15(1), 11-19.
15 <https://doi.org/10.1016/j.tics.2010.10.002>

16 Seymour, B., Daw, N. D., Roiser, J. P., Dayan, P., & Dolan, R. (2012). Serotonin selectively
17 modulates reward value in human decision-making. *J Neurosci*, 32(17), 5833-5842.
18 <https://doi.org/10.1523/JNEUROSCI.0053-12.2012>

19 Shahar, N., Hauser, T. U., Moutoussis, M., Moran, R., Keramati, M., consortium, N., & Dolan, R.
20 J. (2019). Improving the reliability of model-based decision-making estimates in the two-
21 stage decision task with reaction-times and drift-diffusion modeling. *PLoS Comput Biol*,
22 15(2), e1006803. <https://doi.org/10.1371/journal.pcbi.1006803>

1 Sip, K. E., Gonzalez, R., Taylor, S. F., & Stern, E. R. (2017). Increased Loss Aversion in
2 Unmedicated Patients with Obsessive-Compulsive Disorder. *Front Psychiatry*, 8, 309.
3 <https://doi.org/10.3389/fpsyg.2017.00309>

4 Sokol-Hessner, P., Hsu, M., Curley, N. G., Delgado, M. R., Camerer, C. F., & Phelps, E. A. (2009).
5 Thinking like a trader selectively reduces individuals' loss aversion. *Proc Natl Acad Sci U
6 S A*, 106(13), 5035-5040. <https://doi.org/10.1073/pnas.0806761106>

7 Sokol-Hessner, P., & Rutledge, R. B. (2019). The Psychological and Neural Basis of Loss Aversion.
8 *Current Directions in Psychological Science*, 28(1), 20-27.
9 <https://doi.org/10.1177/0963721418806510>

10 Speekenbrink, M., & Konstantinidis, E. (2015). Uncertainty and exploration in a restless bandit
11 problem. *Top Cogn Sci*, 7(2), 351-367. <https://doi.org/10.1111/tops.12145>

12 Stauffer, W. R., Lak, A., & Schultz, W. (2014). Dopamine reward prediction error responses
13 reflect marginal utility. *Current biology*, 24(21), 2491–2500.
14 <https://doi.org/10.1016/j.cub.2014.08.064>

15 Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (Second ed.). MIT
16 Press.

17 Teufel, C., & Fletcher, P. C. (2016). The promises and pitfalls of applying computational models
18 to neurological and psychiatric disorders. *Brain*, 139(Pt 10), 2600-2608.
19 <https://doi.org/10.1093/brain/aww209>

20 Tobler, P. N., Christopoulos, G. I., O'Doherty, J. P., Dolan, R. J., & Schultz, W. (2009). Risk-
21 dependent reward value signal in human prefrontal cortex. *PNAS*, 106(17), 7185–7190.
22 <https://doi.org/10.1073/pnas.0809599106>

1 Tremeau, F., Brady, M., Saccente, E., Moreno, A., Epstein, H., Citrome, L., Malaspina, D., &

2 Javitt, D. (2008). Loss aversion in schizophrenia. *Schizophr Res*, 103(1-3), 121-128.

3 <https://doi.org/10.1016/j.schres.2008.03.027>

4 Tversky, A., & Kahneman, D. (1992). Advances in Prospect Theory: Cumulative Representation

5 of Uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297-323. <https://doi.org/10.1007/Bf00122574>

6

7 Valton, V., Wise, T., & Robinson, O. J. (2020). *Recommendations for Bayesian hierarchical model*

8 *specifications for case-control studies in mental health* Machine Learning for Health

9 (ML4H) at NeurIPS 2020, 34th Conference on Neural Information Processing Systems.

10 arXiv:2011.01725 [cs.CY] Ithaca.

11 Waltmann, M., Schlagenhauf, F., & Deserno, L. (2022). Sufficient reliability of the behavioral and

12 computational readouts of a probabilistic reversal learning task. *Behavior research*

13 *methods*, 10.3758/s13428-021-01739-7.

14 Wang, X. J., & Krystal, J. H. (2014). Computational psychiatry. *Neuron*, 84(3), 638-654.

15 <https://doi.org/10.1016/j.neuron.2014.10.018>

16 Wiecki, T. V., Poland, J., & Frank, M. J. (2015). Model-Based Cognitive Neuroscience Approaches

17 to Computational Psychiatry. *Clinical Psychological Science*, 3(3), 378-399.

18 <https://doi.org/10.1177/2167702614565359>

19 Wilson, R. C., & Collins, A. G. (2019). Ten simple rules for the computational modeling of

20 behavioral data. *Elife*, 8. <https://doi.org/10.7554/eLife.49547>

1 Yi, M. S. K., Steyvers, M., & Lee, M. (2009). Modeling Human Performance in Restless Bandits
2 with Particle Filters. *The Journal of Problem Solving*, 2(2). [https://doi.org/10.7771/1932-
3 6246.1060](https://doi.org/10.7771/1932-6246.1060)