**Supplemental Materials**

**Non-Centered ORL Model Parameterization**

In the main text, we present the centered parameterization of the person-level parameters in the ORL model for clarity; however, to increase efficiency, we implemented the model using non-centered parameterizations. For the non-centered parameterization, person-level parameters are assumed to following separate and independent, standard normal distributions:

|  |  |  |
| --- | --- | --- |
|  | ~ | Equation S1 |

where *zi1* and *zi2* are the *z*-scored person-level parameters for participant *i* on sessions 1 and 2, respectively. To obtain estimates that follow a multivariate normal distribution, we performed the following calculation:

|  |  |  |
| --- | --- | --- |
|  |  | Equation S2 |

where *θi1* and *θi2* are the person-level parameters for participant *i* on sessions 1 and 2, respectively (cf. Equation 7); *μ*1 and *μ*1 are separate and independent group-level means; and is the Cholesky factor of the covariance matrix:

|  |  |  |
| --- | --- | --- |
|  |  | Equation S3 |

where *σθ*,1 and *σθ*,2 are the group-level standard deviations for sessions 1 and 2, respectively, and is the Cholesky factor of the correlation matrix (**R*θ***) between session 1 and session 2 person-level parameters. We reconstructed the correlation matrix **R*θ*** by multiplying by its transpose:

|  |  |  |
| --- | --- | --- |
|  |  | Equation S4 |

**ORL Model Simulations**

Results from fitting the joint ORL model revealed that the group-level estimates of reward learning rate decreased, and perseverance increased across sessions. To illustrate the behavioral changes associated with changes in these two parameters, we simulated data from 1000 hypothetical participants using posterior means for the group-level parameter estimates from session 1 and 2. Session 1 data were simulated using session 1 posterior means for all parameters. Δ*A*+ data were simulated using the session 2 posterior mean for reward learning rate and session 1 posterior means for the other parameters, allowing us to show the behavior changes associated with decreases in reward learning rate while holding the other parameters, importantly perseverance, at their “baseline” levels. Similarly, Δ*βp* data were simulated using the session 2 posterior mean for perseverance and session 1 posterior means for the other parameters, allowing us to show the behavior changes associated with increases in perseverance while holding the other parameters, importantly reward learning rate, at their baseline levels. Finally, Session 2 data were simulated using session 2 posterior means for all parameters, illustrating the behavior changes associated with the combined changes in all parameters, but notably in reward learning rate and perseverance.

Figure S1 shows the results of the simulations. Overall, the results of the simulations are consistent with the behavior changes described in the main text. An increase in reward learning rate is associated with a more rapid decline in playing on a bad deck, specifically Deck B, across trials within-session. A decrease in perseverance is associated with playing more frequently across all decks. Combined, a decrease in reward learning rate and increase in perseverance is associated with playing more frequently on good decks and relatively similar plays on bad decks, on average.

**Figure S1**

*Simulated Play Proportions for A+ & βp Separately & Combined*



*Note.* Simulated proportion of plays across trials for each deck.

**ORL Without Memory Decay**

Because we found low reliability for the memory decay parameter, we refit the ORL without this parameter. To do this, we removed *PSj* from the value-function,

|  |  |  |
| --- | --- | --- |
|  |  | Equation S1 |

where *Vj*(*t* + 1) is the value of playing on deck *j* on the next trial (i.e., *t* + 1), *EVj*(*t* + 1) is the expected outcome value associated with playing or passing on deck *j* in the next trial, *EFj*(*t* + 1) is the expected win frequency of playing or passing on deck *j* in the next trial, *βf* is a free parameter describing sensitivity to win frequency, and *βb* is a free parameter that serves as a bias parameter. The primary difference between this version of the ORL and that presented in the main text is that we removed *PSj* and renamed *βp* to *βb* because *βp* no longer serves as a parameter describing an individual’s tendency to perseverate. Instead, *βb* serves as a bias parameter in which values greater than 0 describes a participant’s tendency to play and values less than 0 describes a participant’s tendency to pass. All other terms are the same as those described in the main text.

Overall, results from the model without memory decaywere similar to that described in the main text. Table S1 shows descriptive statistics and test-retest reliability estimates and Figures S2 and S3 show the mean-level and rank-order stability of ORL parameters, respectively. For the mean-level stability of parameters, reward learning rates (*A*+) decreased, and bias (*βb*) increased, while all other parameters remained relatively the same from session 1 to session 2 (see Table S1 & Figure S1). These findings are qualitatively similar to those of the model with memory decay. For the rank-order stability of parameters, removing memory decay increased reliability for reward learning rates (*A*+) and decreased reliability for punishment learning rates (*A-*), win-frequency sensitivity (*βf*), and bias(*βb*). A difference between these results and those presented in the main text is that the 95% credible interval for the reliability estimate of reward learning rates no longer contains 0. Thus, removing the memory decay parameter improved reliability of reward learning rates.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table S1**  *Descriptive Statistics & Test-Retest Reliability Estimates for Parameters from ORL w/o Memory Decay* | | | | |
| Estimate | *A+* | *A-* | *βf* | *βb* |
| Session 1 *M* [95% CI] | .17 [.13,.21] | .10 [.08,.12] | 2.42 [1.46,3.36] | 0.77 [0.55,0.98] |
| Session 2 *M* [95% CI] | .10 [.07,.13] | .10 [.07,.12] | 2.76 [1.26,4.15] | 1.43 [1.11,1.75] |
| Session 2–1 *M* [95% CI] | **-.07 [-.12,-.02]** | -.003 [-.03,.02] | 0.33 [-1.09,1.80] | **0.67 [0.35,1.00]** |
| *r* [95% CI] | **.67 [.07,1.00]** | **.66 [.28,.99]** | **.59 [.29,.86]** | **.60 [.24,.89]** |

**Figure S2**

*Absolute Stability of Parameters from ORL w/o Memory Decay*



*Note.* Posterior distributions of the group-level ORL parameters from session 1 (left) and session 2 (middle), as well as the difference between session 1 and session 2 estimates (right). Solid red vertical lines represent the posterior means, and the dashed red vertical lines represent the lower and upper bounds of the 95% credible intervals.

**Figure S3**

*Relative Stability of Parameters from ORL w/o Memory Decay*



*Note.* Posterior distributions of the reliability coefficients estimated for each parameter in the joint ORL model. Solid red vertical lines represent the posterior means, horizontal lines represent the lower and upper bounds of the 95% credible intervals, and values to the left and rights sides of each panel represent the % of reliability estimates below and above 0, respectively.

**Session-Specific Correlations**

In the main text, we present correlations between the self-report measures and IGT measures in which measures were averaged across sessions prior to calculating the correlations. Below, we present correlations between IGT measures and self-reports for session 1 (Table S2) and session 2 (Table S3).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table S2**  *Correlations between Self-Report Scores and IGT Measures During Session 1* | | | | | | | | |
|  | Proportion of Plays | |  | ORL Parameters | | | | |
| Self-Report | Good Decks | Bad Decks |  | *A+* | *A-* | *βf* | *βp* | *K* |
| BAS Total | -.17 [-.41,.12] | .09 [-.22,.36] |  | **.26 [.01,.43]** | .03 [-.20,.26] | -.07 [-.28,.14] | -.07 [-.29,.14] | **-.31 [-.42,-.15]** |
| BAS Drive | -.25 [-.51,.06] | .12 [-.16,.43] |  | **.32 [.08,.49]** | -.07 [-.34,.15] | -.15 [-.35,.06] | -.16 [-.43,.11] | **-.37 [-.45,-.22]** |
| BAS Fun | .04 [-.25,.33] | -.11 [-.46,.23] |  | -.07 [-.31,.11] | **.26 [.03,.44]** | -.02 [-.20,.16] | .22 [-.04,.44] | **-.31 [-.50,-.08]** |
| BAS Reward Sensitivity | -.15 [-.38,.09] | .19 [-.06,.43] |  | .35 [.20,.47]\* | -.09 [-.30,.14] | .02 [-.19,.23] | **-.20 [-.34,-.06]** | -.05 [-.25,.10] |
| BIS Total | .21 [-.10,.48] | -.09 [-.33,.19] |  | **-.30 [-.46,-.13]** | .15 [-.07,.35] | .13 [-.06,.31] | .15 [-.14,.36] | .13 [-.17,.38] |
| PANAS PA | -.09 [-.35,.21] | .21 [-.20,.49] |  | .14 [-.03,.30] | .03 [-.18,.21] | **-.28 [-.43,-.11]** | -.03 [-.22,.16] | **-.19 [-.33,-.05]** |
| PANAS NA | -.13 [-.38,.12] | **-.32 [-.55,-.08]** |  | **-.22 [-.33,-.10]** | .22 [-.03,.50] | -.13 [-.31,.04] | -.05 [-.38,.16] | **.32 [.02,.59]** |
| MASQ General Distress Anxiety | .11 [-.14,.37] | -.25 [-.49,.04] |  | **-.32 [-.44,-.16]** | **.29 [.04,.51]** | -.04 [-.24,.14] | **.23 [.09,.35]** | -.03 [-.13,.07] |
| MASQ General Distress Depression | .05 [-.18,.30] | -.16 [-.44,.10] |  | **-.19 [-.33,-.03]** | .20 [-.06,.46] | -.18 [-.38,.03] | **.21 [.09,.35]** | -.09 [-.19,.04] |
| MASQ Anxious Arousal | .11 [-.17,.39] | -.20 [-.48,.25] |  | **-.19 [-.31,-.06]** | **.31 [.05,.52]** | -.01 [-.19,.15] | .21 [-.00,.39] | .03 [-.12,.15] |
| MASQ Anhedonic Depression | .09 [-.24,.38] | -.17 [-.45,.22] |  | **-.25 [-.41,-.09]** | -.08 [-.30,.15] | .06 [-.11,.23] | .05 [-.22,.30] | **.28 [.10,.45]** |
| SHAPS | -.05 [-.26,.17] | -.10 [-.41,.14] |  | .03 [-.33,.24] | .09 [-.04,.21] | -.01 [-.24,.25] | -.00 [-.16,.22] | -.16 [-.39,.06] |
| PROMIS-D | .13 [-.21,.42] | -.10 [-.38,.30] |  | **-.20 [-.35,-.04]** | -.05 [-.26,.17] | -.00 [-.20,.19] | .06 [-.24,.33] | .21 [-.06,.47] |
| *Note.* Correlations with bootstrapped 95% confidence intervals. Confidence intervals that do not overlap with 0 are bolded. | | | | | | | | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table S3**  *Correlations between Self-Report Scores and IGT Measures During Session 2* | | | | | | | | |
|  | Proportion of Plays | |  | ORL Parameters | | | | |
| Self-Report | Good Decks | Bad Decks |  | *A+* | *A-* | *βf* | *βp* | *K* |
| BAS Total | -.02 [-.41,.44] | -.00 [-.28,.35] |  | -.09 [-.34,.16] | .07 [-.14,.27] | **-.24 [-.42,-.06]** | .14 [-.16,.42] | -.16 [-.50,.25] |
| BAS Drive | -.07 [-.55,.57] | .09 [-.27,.52] |  | .01 [-.27,.26] | -.06 [-.34,.20] | **-.29 [-.48,-.06]** | .06 [-.28,.38] | -.07 [-.29,.23] |
| BAS Fun | .18 [-.21,.52] | .05 [-.32,.40] |  | -.16 [-.43,.10] | **.21 [.05,.35]** | -.12 [-.33,.14] | **.31 [.06,.55]** | -.18 [-.50,.24] |
| BAS Reward Sensitivity | -.19 [-.45,.11] | -.20 [-.44,.06] |  | -.08 [-.27,.11] | .01 [-.17,.22] | -.09 [-.29,.12] | -.09 [-.30,.16] | -.11 [-.31,.15] |
| BIS Total | -.06 [-.50,.31] | -.06 [-.44,.37] |  | -.06 [-.26,.16] | -.11 [-.28,.07] | .13 [-.11,.35] | -.11 [-.36,.13] | .11 [-.06,.23] |
| PANAS PA | .16 [-.12,.40] | .08 [-.24,.40] |  | **.26 [.04,.43]** | -.02 [-.22,.18] | .04 [-.19,.28] | .13 [-.04,.31] | .10 [-.02,.24] |
| PANAS NA | -.08 [-.40,.23] | -.14 [-.43,.19] |  | -.05 [-.26,.16] | .19 [-.04,.41] | -.06 [-.26,.12] | -.02 [-.27,.21] | **.33 [.06,.52]** |
| SHAPS | -.14 [-.36,.17] | .04 [-.29,.40] |  | -.07 [-.38,.24] | .06 [-.09,.20] | **-.43 [-.59,-.23]** | -.10 [-.25,.06] | .02 [-.14,.15] |
| *Note.* Correlations with bootstrapped 95% confidence intervals. Confidence intervals that do not overlap with 0 are bolded. | | | | | | | | |