Supplementary Information

A. Snapshot of the Repeated Trust Game as Seen by Participants

Figure S1 shows a screenshot of the repeated Trust Game at the moment the participant is required to make a decision of how much to send back to the Investor.

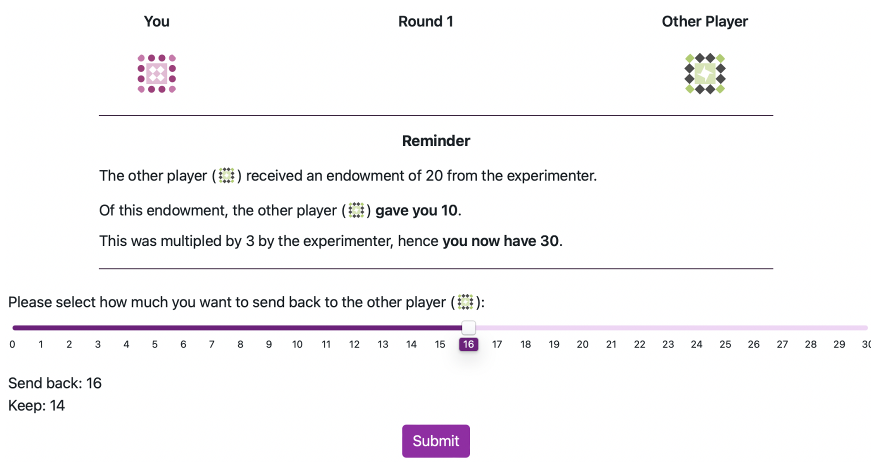

Figure S1: Screenshot of the RTG as seen by participants at the decision phase.

B. Hidden Markov Model Used to Simulate the Investor's Actions

The HMM assumes that the probability of each investment $I_t = 0, \ldots, 20$, at each trial t, conditional on the current state of the investor S_t, is dependent on an underlying normal distribution with mean μ_s and standard deviation σ_s. The probability of each discrete investment was determined from the cumulative normal distribution Φ, computing the probability of a Normal variate falling between the midway points of the response options. As responses were bounded at 0 and 20, we normalized these probabilities further by taking the endpoints into account. For instance, the probability of an investment $I_t = 2$ is defined as:

$$P(I_t = 2|S_t = s) = \frac{\Phi(2.5|\mu_s, \sigma_s) - \Phi(1.5|\mu_s, \sigma_s)}{\Phi(20.5|\mu_s, \sigma_s) - \Phi(-0.5|\mu_s, \sigma_s)}$$

Note that the denominator truncates the distribution between 0 and 20. To estimate the transition probability between states for the investor, a multinomial logistic regression model was fitted to the investor's data such as:

$$P(S_{t+1} = s'|S_t = s, X_t = x) = \frac{\exp(\beta_{0,s,s'} + \beta_{1,s,s'}x)}{\sum_{s''} \exp(\beta_{0,s,s''} + \beta_{1,s,s''}x)}$$

where $X_t = R_t - I_t$ is the net return to the investor with R_t the amount returned by the trustee and I_t is the Investment sent.

The advantages of this approach are that it does not require any a priori assumptions about the model features. The number of states, the policy conditional on the state, and the transition function between states can all be determined in a purely data-driven way. These HMMs can in turn be used to simulate a human-like agent playing the trust game. This agent may transition to a new state depending on the other player's actions and adopt a policy reflecting its state, thus simulating changes in emotional dispositions of human players during a repeated game. When the investor gains from the interaction, they become more likely to transition to a state where their policy is more "trusting" with generally higher investments. However, faced with losses, the investor is more likely to transition to a more cautious policy with generally lower investments. The policies and the transitions between states are sufficient to build an agent that reflects this type of adaptive behavior and reacts to the trustee's action choices in a way that mimics a human player.

We estimated a three-state model for the investor's behavior, using maximum likelihood estimation via the Expectation-Maximization algorithm as implemented in the depmixS4 package for R (Visser & Speekenbrink, 2021). The model was estimated using investments from existing datasets of human dyads playing 10 rounds of the RTG with the same trustee. The dataset consisted of a total of 381 games from two data sources: First, a total of 93 repeated trust games with healthy investors and a mix of healthy trustees and trustees diagnosed with Borderline Personality Disorder (BPD) (King-Casas et al., 2008). The second source was from data collected as part of a project investigating social exchanges in BPD and antisocial personality disorder reported elsewhere (Euler et al., 2021; Huang et al., 2020; Rifkin-Zybutz et al., 2021) and consists of 288 games. In both datasets, the investor on which we modeled the HMM's strategy was always selected from a healthy population and the trustees were a mix of healthy participants and those with personality disorders, allowing for a diversified interaction behavior.

C. Mixed-effects Models for Participant Returns

We fit a linear mixed effects model to participant returns as a proportion of the multiplied investment received as described below. The results of the model are presented in Table 1.

$$
\begin{aligned}
\mathrm{R}_{ij} = {} & \beta_0 + \beta_1 \text{ Phase}_i + \beta_2 \text{ Condition}_i + \beta_3 \text{ Investment}_i + \beta_4 \text{ RS}_i + \\
& \beta_5(\text{Phase} \times \text{Condition})_i + \beta_6(\text{Phase} \times \text{Investment})_i + \beta_7(\text{Phase} \times \text{RS})_i + \\
& \beta_8(\text{Condition} \times \text{Investment})_i + \beta_9(\text{Condition} \times \text{RS})_i + \beta_{10}(\text{Investment} \times \text{RS})_i + \\
& \beta_{11}(\text{Phase} \times \text{Condition} \times \text{Investment})_i + \beta_{12}(\text{Phase} \times \text{Condition} \times \text{RS})_i + \\
& \beta_{13}(\text{Phase} \times \text{Investment} \times \text{RS})_i + \beta_{14}(\text{Condition} \times \text{Investment} \times \text{RS})_i + \\
& \beta_{15}(\text{Phase} \times \text{Condition} \times \text{Investment} \times \text{RS})_i + \\
& b_{0j} + b_{1j} \text{ (Phase)}_i + \epsilon_{ij}
\end{aligned}
$$

where:

- R_{ij}: percentage of tripled investment returned to investor for participant j in observation i
- β_0: intercept
- β_1 to β_4: main effects of Phase (RTG pre vs. post-manipulation), Condition (manipulation vs. control), Investment, and RS (High vs Low RS), respectively
- β_5 to β_{10}: interaction effects between each pair of the four factors, showing how the relationship between one factor and the return percentage changes depending on the level of another factor
- β_{11} to β_{14}: three-way interaction effects among the four factors, indicating how the interaction between two factors is further modified by the third factor
- β_{15}: four-way interaction effect between Phase, Condition, Investment, and RS, describing how the interaction among three factors is modified by the fourth factor
- b_{0j}: player-wise random intercept for player j
- b_{1j}: player-wise random slope for Phase for player j
- ϵ_{ij}: error term for player j in observation i

Table 1: Summary of Mixed-Effects Model of participant returns across all rounds

Term	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.43	0.01	200.83	49.11	< .001
Phase	0.01	0.00	199.61	2.42	.017
Condition	0.00	0.01	200.83	-0.37	.714
Investment	0.04	0.00	5928.40	18.06	< .001
RS_group	0.01	0.01	200.83	1.44	.151
Phase:Condition	-0.01	0.00	199.61	-2.10	.037
Phase:Investment	0.00	0.00	5839.71	-0.84	.399
Condition:Investment	0.01	0.00	5928.40	3.68	< .001
Phase:RS_group	0.01	0.00	199.61	1.55	.123
Condition:RS_group	0.01	0.01	200.83	0.65	.515
Investment:RS_group	0.00	0.00	5928.40	0.42	.678
Phase:Condition:Investment	0.00	0.00	5839.71	-0.31	.758
Phase:Condition:RS_group	0.00	0.00	199.61	-0.57	.571
Phase:Investment:RS_group	-0.01	0.00	5839.71	-2.86	.004
Condition:Investment:RS_group	0.00	0.00	5928.40	-0.87	.382
Phase:Condition:Investment:RS_group	0.01	0.00	5839.71	3.10	.002

D. Mixed-effects Models for HMM Investments

We fit a linear mixed effects model to the investments sent by the HMM with Phase, Condition and RS groups as fixed effects, and player-wise intercepts. The results of the model are presented in Table 2.

Table 2: Summary of Mixed-Effects Model of HMM investments across all rounds

Term	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	9.87	0.16	202	60.60	< .001
Phase	0.04	0.14	202	0.25	.803
Condition	0.15	0.16	202	0.90	.367
RS_group	0.21	0.16	202	1.27	.205
Phase:Condition	0.01	0.14	202	0.04	.970
Phase:RS_group	0.06	0.14	202	0.43	.668
Condition:RS_group	0.00	0.16	202	0.01	.993
Phase:Condition:RS_group	-0.12	0.14	202	-0.84	.400

E. Exploitation Diagnostic Model

We fit a mixed logistic regression model to predict whether a participant's return constituted exploitation (returning less than one-third of the tripled investment). The model included Phase, Condition, Investment (scaled), and RS group as fixed effects, with player-wise random intercepts. The results are presented in Table 3.

$$\text{logit}(P(\text{Exploit}_{ij} = 1)) = \beta_0 + \beta_1 \text{ Phase}_i + \beta_2 \text{ Condition}_i + \beta_3 (\text{Phase} \times \text{Condition})_i + \beta_4 \text{ Investment}_i + \beta_5 \text{ RS}_i + b_{0j}$$

where:

- Exploit_{ij}: binary outcome indicating whether participant j returned less than one-third of the tripled investment in observation i
- β_0: intercept (log-odds)
- β_1 to β_3: fixed effects of Phase, Condition, and their interaction
- β_4, β_5: fixed effects of Investment (scaled) and RS group
- b_{0j}: player-wise random intercept for player j

Table 3: Summary of Mixed Logistic Regression for Exploitation Probability

Term	Estimate	Std. Error	z value	Pr(>|z|)
Intercept	-2.70	0.26	-10.18	< .001
Phase	0.37	0.12	3.18	.001
Condition	-0.49	0.31	-1.57	.116
Investment	-0.96	0.05	-18.72	< .001
RS_group	0.24	0.30	0.82	.410
Phase:Condition	0.18	0.17	1.06	.291

F. Pre-Defection Returns Model

To examine whether differences between conditions emerged before the pre-programmed low investment, we fit the same four-way mixed effects model as in Section C, restricted to rounds prior to the low investment (pre-phase: rounds 1–11; post-phase: rounds 1–12). The results are presented in Table 4.

$$\begin{aligned} \text{R}_{ij} = & \beta_0 + \beta_1 \text{ Phase}_i + \beta_2 \text{ Condition}_i + \beta_3 \text{ Investment}_i + \beta_4 \text{ RS}_i + \\ & \beta_5 (\text{Phase} \times \text{Condition})_i + \beta_6 (\text{Phase} \times \text{Investment})_i + \beta_7 (\text{Phase} \times \text{RS})_i + \\ & \beta_8 (\text{Condition} \times \text{Investment})_i + \beta_9 (\text{Condition} \times \text{RS})_i + \beta_{10} (\text{Investment} \times \text{RS})_i + \\ & \beta_{11} (\text{Phase} \times \text{Condition} \times \text{Investment})_i + \beta_{12} (\text{Phase} \times \text{Condition} \times \text{RS})_i + \\ & \beta_{13} (\text{Phase} \times \text{Investment} \times \text{RS})_i + \beta_{14} (\text{Condition} \times \text{Investment} \times \text{RS})_i + \\ & \beta_{15} (\text{Phase} \times \text{Condition} \times \text{Investment} \times \text{RS})_i + \\ & b_{0j} + b_{1j} \ (\text{Phase})_i + \epsilon_{ij} \end{aligned}$$

All variables are defined as in Section C. The analysis is restricted to rounds before the pre-programmed low investment to test whether contrast effects were present prior to any trust violation.

G. Event Study Model

We fit a linear mixed effects model to participant returns within a 5-round window centered on the pre-programmed low investment ($t-2$ to $t+2$). The time_point factor uses custom contrasts to test the Drop

Table 4: Summary of Mixed-Effects Model of Participant Returns (Pre-Defection Rounds Only)

Term	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.43	0.01	200.62	49.85	< .001
Phase	0.01	0.00	199.84	1.95	.053
Condition	0.00	0.01	200.62	-0.20	.840
Investment	0.03	0.00	4513.25	10.84	< .001
RS_group	0.01	0.01	200.62	1.53	.127
Phase:Condition	-0.01	0.00	199.84	-2.20	.029
Phase:Investment	0.00	0.00	4454.13	-0.46	.648
Condition:Investment	0.00	0.00	4513.25	1.35	.178
Phase:RS_group	0.01	0.00	199.84	1.97	.050
Condition:RS_group	0.01	0.01	200.62	0.92	.357
Investment:RS_group	0.00	0.00	4513.25	-0.07	.947
Phase:Condition:Investment	0.00	0.00	4454.13	0.94	.346
Phase:Condition:RS_group	0.00	0.00	199.84	-0.74	.462
Phase:Investment:RS_group	-0.01	0.00	4454.13	-3.72	< .001
Condition:Investment:RS_group	0.00	0.00	4513.25	-1.51	.132
Phase:Condition:Investment:RS_group	0.01	0.00	4454.13	2.67	.008

(change at defection relative to $t-1$) and Recovery (change at $t+1$ relative to defection). The results are presented in Table 5.

$$\begin{aligned} \mathrm{R}_{ij} = &\beta_0 + \beta_1 \text{ Phase}_i + \beta_2 \text{ Condition}_i + \beta_3 \text{ TimePoint}_i + \beta_4 \text{ RS}_i + \\ &\text{all two-way, three-way, and four-way interactions} + \\ &b_{0j} + \epsilon_{ij} \end{aligned}$$

where:

- R_{ij}: percentage of tripled investment returned for participant j in observation i
- TimePoint: factor with 5 levels ($t-2$, $t-1$, Defection, $t+1$, $t+2$) centered on the pre-programmed low investment round
- b_{0j}: player-wise random intercept for player j

All other variables are defined as in Section C.

H. Exposure Phase Returns Model

We fit a linear mixed effects model to participant returns during the three exposure phase games (expo1, expo2, expo3). The model includes Condition, Investment (scaled), and RS group as fixed effects with their three-way interaction, and player-wise random intercepts. The results are presented in Table 6.

$$\begin{aligned} \mathrm{R}_{ij} = &\beta_0 + \beta_1 \text{ Condition}_i + \beta_2 \text{ Investment}_i + \beta_3 \text{ RS}_i + \\ &\beta_4(\text{Condition} \times \text{Investment})_i + \beta_5(\text{Condition} \times \text{RS})_i + \beta_6(\text{Investment} \times \text{RS})_i + \\ &\beta_7(\text{Condition} \times \text{Investment} \times \text{RS})_i + \\ &b_{0j} + \epsilon_{ij} \end{aligned}$$

where all variables are defined as in Section C, but Phase is absent because the exposure phase is a single time point.

Table 5: Summary of Mixed-Effects Model for Event Study Analysis (Drop and Recovery)

Term	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.41	0.01	201.10	37.70	< .001
Phase	0.01	0.00	1773.30	3.86	< .001
Condition	0.00	0.01	201.10	-0.39	.699
TimePoint_1	0.02	0.01	1773.27	2.67	.008
TimePoint_2	0.02	0.01	1773.71	2.36	.018
TimePoint_3	-0.07	0.01	1772.40	-9.30	< .001
TimePoint_4	0.02	0.01	1773.50	2.08	.037
RS_group	0.02	0.01	201.10	1.39	.165
Phase:Condition	-0.01	0.00	1773.30	-1.68	.094
Phase:TimePoint_1	0.00	0.01	1773.20	0.36	.721
Phase:TimePoint_2	0.00	0.01	1773.83	0.36	.717
Phase:TimePoint_3	-0.01	0.01	1772.41	-1.10	.272
Phase:TimePoint_4	0.00	0.01	1773.61	0.09	.926
Condition:TimePoint_1	0.01	0.01	1773.27	0.83	.407
Condition:TimePoint_2	0.01	0.01	1773.71	1.64	.102
Condition:TimePoint_3	-0.01	0.01	1772.40	-1.80	.073
Condition:TimePoint_4	0.00	0.01	1773.50	0.48	.634
Phase:RS_group	0.01	0.00	1773.30	1.37	.170
Condition:RS_group	0.01	0.01	201.10	0.58	.563
TimePoint_1:RS_group	0.01	0.01	1773.27	0.79	.431
TimePoint_2:RS_group	0.00	0.01	1773.71	0.14	.885
TimePoint_3:RS_group	-0.01	0.01	1772.40	-0.81	.417
TimePoint_4:RS_group	0.00	0.01	1773.50	0.26	.796
Phase:Condition:TimePoint_1	0.00	0.01	1773.20	0.36	.720
Phase:Condition:TimePoint_2	-0.01	0.01	1773.83	-1.02	.306
Phase:Condition:TimePoint_3	0.01	0.01	1772.41	1.05	.294
Phase:Condition:TimePoint_4	-0.01	0.01	1773.61	-0.98	.329
Phase:Condition:RS_group	0.00	0.00	1773.30	-0.46	.642
Phase:TimePoint_1:RS_group	0.01	0.01	1773.20	1.06	.288
Phase:TimePoint_2:RS_group	0.01	0.01	1773.83	1.15	.249
Phase:TimePoint_3:RS_group	-0.01	0.01	1772.41	-1.63	.103
Phase:TimePoint_4:RS_group	0.00	0.01	1773.61	-0.44	.657
Condition:TimePoint_1:RS_group	0.00	0.01	1773.27	0.54	.590
Condition:TimePoint_2:RS_group	0.02	0.01	1773.71	2.00	.046
Condition:TimePoint_3:RS_group	0.00	0.01	1772.40	0.53	.598
Condition:TimePoint_4:RS_group	-0.02	0.01	1773.50	-2.24	.026
Phase:Condition:TimePoint_1:RS_group	0.00	0.01	1773.20	-0.16	.876
Phase:Condition:TimePoint_2:RS_group	0.00	0.01	1773.83	-0.62	.534
Phase:Condition:TimePoint_3:RS_group	0.00	0.01	1772.41	-0.49	.621
Phase:Condition:TimePoint_4:RS_group	0.01	0.01	1773.61	0.84	.403

We additionally fit a linear mixed effects model to the HMM investments during the exposure phase with Condition and RS group as fixed effects, restricted to rounds 5 and above to allow sufficient time for the HMM's adaptive strategy to differentiate between conditions. The results are presented in Table 7.

I. Four-Way Interaction Slopes Analysis

To decompose the significant four-way interaction (Phase $\times$ Condition $\times$ Investment $\times$ RS), we used `emtrends()` to estimate the marginal slope of investment on returns for each combination of Phase, Condition, and RS group. Custom contrasts then tested whether the investment slope changed from pre to post within each condition and RS group. The slopes and contrasts are presented in Tables 8 and 9.

Table 6: Summary of Mixed-Effects Model of Participant Returns During Exposure Phase

Term	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.44	0.01	202.51	46.62	$< .001$
Condition	0.00	0.01	202.51	0.41	.683
Investment	0.04	0.00	4117.02	15.27	$< .001$
RS_group	0.01	0.01	202.51	1.23	.222
Condition:Investment	0.02	0.00	4117.02	6.78	$< .001$
Condition:RS_group	0.01	0.01	202.51	1.43	.155
Investment:RS_group	0.00	0.00	4117.02	-1.21	.228
Condition:Investment:RS_group	-0.01	0.00	4117.02	-2.05	.040

Table 7: Summary of Mixed-Effects Model of HMM Investments During Exposure Phase

Term	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	9.76	0.15	202	66.51	$< .001$
Condition	-2.06	0.15	202	-14.06	$< .001$
RS_group	0.10	0.15	202	0.65	.515
Condition:RS_group	0.05	0.15	202	0.36	.716

Table 8: Estimated Marginal Slopes of Investment on Returns by Phase, Condition, and RS Group

phase.f	condition.f	high_RS	inv_scaled.trend	SE	df	lower.CL	upper.CL
pre	Control	low RS	0.043	0.006	5604.447	0.031	0.054
post	Control	low RS	0.047	0.006	5728.831	0.035	0.058
pre	Manipulation	low RS	0.019	0.006	5786.137	0.008	0.031
post	Manipulation	low RS	0.046	0.006	5813.694	0.035	0.058
pre	Control	high RS	0.044	0.006	5625.162	0.032	0.055
post	Control	high RS	0.049	0.006	5828.678	0.038	0.061
pre	Manipulation	high RS	0.039	0.006	5749.580	0.027	0.051
post	Manipulation	high RS	0.016	0.006	5836.386	0.004	0.028

Table 9: Pre-to-Post Slope Change Contrasts (Sidak-adjusted)

contrast	estimate	SE	df	t.ratio	p.value
ManipLowRS	0.027	0.008	5859.003	3.260	.003
ManipHighRS	-0.023	0.008	5865.216	-2.666	.023
HighVsLowRS	-0.050	0.012	5862.765	-4.186	$< .001$

J. Player Rating Models

We fit three separate linear mixed effects models to participants' ratings of their co-players: perceived cooperativeness, perceived forgiveness, and willingness to play again. Each model included Game Phase (pre, expo1, expo2, expo3, post), Condition, and RS group as fixed effects with their three-way interaction, and player-wise random intercepts. The results are presented in Tables 10, 11, and 12.

$$\text{Rating}_{ij} = \beta_0 + \beta_1 \text{ GamePhase}_i + \beta_2 \text{ Condition}_i + \beta_3 \text{ RS}_i + \text{all interactions} + b_{0j} + \epsilon_{ij}$$

where:

- Rating$_{ij}$: the rating (cooperativeness, forgiveness, or willingness to play again) by participant j at phase i
- GamePhase: factor with 5 levels (pre, expo1, expo2, expo3, post)
- All other variables are defined as in Section C

Table 10: Summary of Mixed-Effects Model for Perceived Cooperativeness Ratings

Term	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	6.31	0.11	202	57.69	< .001
Phase_1	0.16	0.11	808	1.41	.158
Phase_2	0.24	0.11	808	2.11	.035
Phase_3	0.06	0.11	808	0.49	.622
Phase_4	-0.08	0.11	808	-0.70	.484
Condition	-0.03	0.11	202	-0.26	.794
RS_group	0.23	0.11	202	2.09	.038
Phase_1:Condition	0.34	0.11	808	2.98	.003
Phase_2:Condition	-0.33	0.11	808	-2.94	.003
Phase_3:Condition	-0.02	0.11	808	-0.18	.858
Phase_4:Condition	-0.25	0.11	808	-2.24	.026
Phase_1:RS_group	0.04	0.11	808	0.32	.752
Phase_2:RS_group	-0.07	0.11	808	-0.60	.546
Phase_3:RS_group	0.08	0.11	808	0.70	.484
Phase_4:RS_group	-0.13	0.11	808	-1.11	.269
Condition:RS_group	-0.01	0.11	202	-0.09	.924
Phase_1:Condition:RS_group	-0.26	0.11	808	-2.31	.021
Phase_2:Condition:RS_group	-0.06	0.11	808	-0.50	.619
Phase_3:Condition:RS_group	0.18	0.11	808	1.55	.122
Phase_4:Condition:RS_group	0.09	0.11	808	0.78	.434

Table 11: Summary of Mixed-Effects Model for Perceived Forgiveness Ratings

Term	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	5.88	0.11	202	53.04	< .001
Phase_1	0.21	0.12	808	1.80	.073
Phase_2	0.13	0.12	808	1.12	.265
Phase_3	0.13	0.12	808	1.09	.274
Phase_4	0.09	0.12	808	0.73	.465
Condition	-0.05	0.11	202	-0.45	.650
RS_group	0.13	0.11	202	1.13	.260
Phase_1:Condition	0.13	0.12	808	1.05	.292
Phase_2:Condition	-0.28	0.12	808	-2.35	.019
Phase_3:Condition	0.02	0.12	808	0.19	.847
Phase_4:Condition	-0.23	0.12	808	-1.89	.059
Phase_1:RS_group	-0.08	0.12	808	-0.70	.484
Phase_2:RS_group	-0.06	0.12	808	-0.53	.596
Phase_3:RS_group	0.17	0.12	808	1.41	.158
Phase_4:RS_group	-0.10	0.12	808	-0.84	.401
Condition:RS_group	0.02	0.11	202	0.17	.864
Phase_1:Condition:RS_group	-0.22	0.12	808	-1.80	.072
Phase_2:Condition:RS_group	-0.05	0.12	808	-0.42	.678
Phase_3:Condition:RS_group	0.12	0.12	808	0.97	.332
Phase_4:Condition:RS_group	0.07	0.12	808	0.61	.541

Table 12: Summary of Mixed-Effects Model for Willingness to Play Again Ratings

Term	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	6.24	0.12	202	49.91	< .001
Phase_1	0.58	0.14	808	4.18	< .001
Phase_2	0.16	0.14	808	1.16	.246
Phase_3	-0.15	0.14	808	-1.11	.268
Phase_4	-0.06	0.14	808	-0.41	.679
Condition	-0.13	0.12	202	-1.01	.316
RS_group	0.27	0.12	202	2.18	.030
Phase_1:Condition	0.21	0.14	808	1.50	.135
Phase_2:Condition	-0.47	0.14	808	-3.34	< .001
Phase_3:Condition	0.02	0.14	808	0.14	.890
Phase_4:Condition	-0.15	0.14	808	-1.08	.281
Phase_1:RS_group	0.05	0.14	808	0.34	.731
Phase_2:RS_group	-0.06	0.14	808	-0.45	.655
Phase_3:RS_group	0.02	0.14	808	0.16	.874
Phase_4:RS_group	-0.19	0.14	808	-1.38	.169
Condition:RS_group	-0.08	0.12	202	-0.64	.526
Phase_1:Condition:RS_group	-0.22	0.14	808	-1.59	.113
Phase_2:Condition:RS_group	0.04	0.14	808	0.30	.762
Phase_3:Condition:RS_group	0.20	0.14	808	1.46	.144
Phase_4:Condition:RS_group	0.06	0.14	808	0.41	.684

References

Euler, S., Nolte, T., Constantinou, M., Griem, J., Montague, P. R., Fonagy, P., & Personality and Mood Disorders Research Network. (2021). Interpersonal Problems in Borderline Personality Disorder: Associations With Mentalizing, Emotion Regulation, and Impulsiveness. *Journal of Personality Disorders*, *35*(2), 177–193. https://doi.org/10.1521/pedi_2019_33_427

Huang, Y. L., Fonagy, P., Feigenbaum, J., Montague, P. R., Nolte, T., & and London Personality and Mood Disorder Research Consortium. (2020). Multidirectional Pathways between Attachment, Mentalizing, and Posttraumatic Stress Symptomatology in the Context of Childhood Trauma. *Psychopathology*, *53*(1), 48–58. https://doi.org/10.1159/000506406

King-Casas, B., Sharp, C., Lomax-Bream, L., Lohrenz, T., Fonagy, P., & Montague, P. R. (2008). The Rupture and Repair of Cooperation in Borderline Personality Disorder. *Science*, *321*(5890), 806–810. https://doi.org/10.1126/science.1156902

Rifkin-Zybutz, R. P., Moran, P., Nolte, T., Feigenbaum, J., King-Casas, B., Fonagy, P., & Montague, R. P. (2021). Impaired mentalizing in depression and the effects of borderline personality disorder on this relationship. *Borderline Personality Disorder and Emotion Dysregulation*, *8*(1), 15. https://doi.org/10.1186/s40479-021-00153-x

Visser, I., & Speekenbrink, M. (2021). *depmixS4: Dependent Mixture Models - Hidden Markov Models of GLMs and Other Distributions in S4.*