Full Report 2023-06-19

Samuel Thelaus

2023-06-19

Initial data checks

One participant had fewer than 630 total trials (total of 455 trials remained in file) and was missing columns for the final 3 MCQ questions (4-6). However, this participant was still included in analysis.

Descriptive statistics

This section goes through the tests specified in the pre-registration report and specifies observed effect sizes in the sample so far.

There was a total of 199 participants after excluding participants. The mean age of participants was 21.6 (range 18-30) years old. The gender distribution was Female = 78, Male = 120, Other = 1. The distribution of majors for participants was STEM = 67, Business = 128, Humanities = 2, Other = 2.

Table 1: Participant understanding of basic aspects of the task. “MCQ”: The table reports the percentage of correct replies to the 6 MCQ questions. “Pre-/Post-task quiz”: MCQ about the probability of a winning bet in a yellow/blue background session. The question is asked twice: just before the beginning of the task (“Pre”), and immediately after the task is completed (“Post”). “Optimal”: Participants who play optimally, i.e., they pass in all trials of all yellow background sessions. “Cravers”: Participants who bet at least twice across yellow background sessions during the experimental task (this is the definition used in Test 4; see also Table S3).

Participant type	MCQ	Post-game quiz	Total
Optimal	100%	-	100%
Cravers	100%	99%	100%
Total	100%	99%	100%

Table 2: Participant report on whether they have a strategy of play just before they start performing the task.

	No strategy	Not quite sure	Quite confident	Think it is right
n	1	29	120	49

Table 3: Fraction of cravers in the experiment. “At least twice in yellow”: percentage of participants who bet at least twice across yellow background sessions during the experimental task (this is the definition of “craver” used in Test 4). “Bet in yellow”: percentage of participants who bet once or more in the yellow background sessions during the experimental task (reported here for reference).

Craving definition	Test	Control	Total
Bet in yellow	53/99 = 54%	36/100 = 36%	89/199 = 45%
At least twice in yellow	44/99 = 44%	33/100 = 33%	77/199 = 39%

Table 4: Betting rate of all participants, in blue and yellow background sessions.

	Blue	Yellow
n	199.000	199.000
mean	0.966	0.076
median	1.000	0.000
sd	0.091	0.142
min	0.500	0.000
max	1.000	0.742
skew	-3.537	2.369

Test 1

Paired t-test to compare participant betting rate in the high-reward vs. low-reward sessions in the yellow background sessions, one-sided; H0: Betting rate is higher or equal in the low-reward sessions. We expect to be able to reject H0 at a 5% significance level (Prediction 1).

Participant-level test ($N = 199$).

Data was skewed at 1.611 (ideal values within [-1, 1]) and a Shapiro-Wilks test showed non-normality. Data were Box-Cox transformed with lambda = -18. This reduced skew to 0.816 (acceptable value), but SW test still showed non-normality.

Paired t-test using Box-Cox transformed data showed higher betting rate in high reward sessions ($t(198) = -7.201$, $p < .001$). Non-parametric paired Wilcoxon test using non-transformed data showed qualitatively same results ($V = 242.5$ $p < .001$).

Test 2

Logistic mixed effects model predicting betting, with independent variables: potential reward value in the session, uncertainty level in the session, a dummy for treatment (reference: treatment C), a dummy for session colour (reference: blue), the interaction of treatment and session colour, and several control variables (which include the previous decision, to control for choice inertia/stickiness).

Table 5: Logistic mixed model predicting betting (0: skip, 1: bet) for all participants in all trials. Variables are reward value (0: low, 1: high), uncertainty (0: low, 1: high), treatment (0: control, 1: test), session color (0: blue, 1: yellow), previous choice (0: skip, 1: bet), and the interaction between session color and treatment. Controls not shown were age, gender, major, and sequence number.

Variable	m_1
(Intercept)	2.91 (0.269)***
reward_valueHigh	1.556 (0.042)***
uncertaintyHigh	0.145 (0.036)***
treatmentTest	-0.822 (0.227)***
colorYellow	-8.493 (0.113)***
previous_choice	1.913 (0.039)***
sequence_number	-0.047 (0.018)*
treatmentTest:colorYellow	1.476 (0.13)***
R^2	0.836
N	119096

Note:

*p < .05, **p < .01, ***p < .001

Test 3

Logistic mixed effects model predicting betting, with independent variables: potential reward value in the session, uncertainty level in the session, a dummy for treatment (reference: treatment C), a dummy for session colour (reference: blue), the previous decision, and reward exposure, which is the discounted sum of prior rewards as defined in our computational model (more below). The main variable of interest is the reward exposure variable. We expect it to be significantly positive (Prediction 2).

Table 6: Logistic mixed model predicting betting (0: skip, 1: bet) for all participants in all trials. Variables are reward value (0: low, 1: high), uncertainty (0: low, 1: high), treatment (0: control, 1: test), session color (0: blue, 1: yellow), previous choice (0: skip, 1: bet), reward exposure, and the interaction between session color and treatment. Controls not shown were age, gender, major, and sequence number.

Variable	m_1
(Intercept)	2.905 (0.268)***
reward_valueHigh	1.557 (0.042)***
uncertaintyHigh	0.147 (0.037)***
treatmentTest	-0.817 (0.227)***
colorYellow	-8.517 (0.117)***
previous_choice	1.915 (0.039)***
reward_exposure	-0.02 (0.024)
sequence_number	-0.047 (0.018)*
treatmentTest:colorYellow	1.508 (0.135)***
Nakagawa R^2	0.836
N	119096

Note:

*p < .05, **p < .01, ***p < .001

Next we show the VIF for the model with previous choice and reward exposure. VIF at 1 suggests there is no multicollinearity and VIF >10 suggests problematic collinearity.

Table 7: VIF for variables in test 3.

Variables	VIF
reward_value	1.003054
uncertainty	1.007319
treatment	1.028250
color	2.068245
age	1.029062
gender	1.015811
major	1.019029
sequence_number	1.001179
previous_choice	2.100844
reward_exposure	1.123431

Stepwise model-building

This section shows the process of building the models from the theoretical version (in preregistration) to the model shown in the report. Models are built with incrementally more variables and each model is compared using R^2.

Table 8: Stepwise comparison of models being built from theoretical base (1) to best fitting model (4). Noteworthy is that model (3) is the model from test 2.

Variable	(1)	(2)	(3)	(4)
(Intercept)	4.616 (0.309)***	4.624 (0.313)***	2.911 (0.268)***	2.905 (0.269)***
reward_valueHigh	1.82 (0.041)***	1.819 (0.041)***	1.556 (0.042)***	1.557 (0.042)***
uncertaintyHigh	0.156 (0.034)***	0.156 (0.034)***	0.145 (0.036)***	0.147 (0.037)***
treatmentTest	-0.885 (0.263)**	-0.882 (0.265)**	-0.822 (0.227)***	-0.817 (0.227)***
colorYellow	-9.991 (0.113)***	-9.991 (0.114)***	-8.492 (0.113)***	-8.517 (0.117)***
treatmentTest:colorYellow	1.881 (0.13)***	1.877 (0.131)***	1.476 (0.129)***	1.508 (0.135)***
sequence_number	-	-0.051 (0.017)**	-0.047 (0.018)*	-0.047 (0.018)*
previous_choice	-	-	1.913 (0.039)***	1.916 (0.039)***
reward_exposure	-	-	-	-0.02 (0.024)
-	-	-	-	-
R^2	0.84	0.84	0.836	0.836
N	119096	119096	119096	119096

Note:

*p < .05, **p < .01, **p < .001

Test 4

Linear regression using as dependent variable participant betting rate in yellow (proxy for craving). We will also run a logistic regression using as dependent variable participant type (0: non craver; 1: craver). A craver is defined as a participant who chooses to bet at least twice across the yellow background sessions during the task.†† The independent variables will include a dummy code for treatment (reference: treatment C) and several control variables which include participant accuracy at the post-yellow-sessions lotteries (to control for the aforementioned miswanting aspect of behaviour). Our main variable of interest is the treatment variable: we expect a significantly positive coefficient on that variable (Prediction 3).

Linear model showed following F-statistic: $F(7, 191) = 2.75$, $p = 0.01$, $R^2 = 0.058$.

Table 9: Linear model predicting betting rates in yellow for all participants. Main variable is the treatment effect (0: control, 1: test). Controls not shown were age, gender, major, and sequence number.

Variable	m_1
(Intercept)	0.057 (0.024)*
treatmentTest	0.041 (0.02)*
R^2	0.058
N	199

Note:

*p < .05, **p < .01, ***p < .001

Table 10: Logistic model predicting whether a participant is a craver (bet at least twice in yellow). Main variable is the treatment effect (0: control, 1: test). Controls not shown were age, gender, major, and sequence number.

Variable	m_1
(Intercept)	-0.446 (0.36)
treatmentTest	0.543 (0.308)
Nakagawa R^2	0.067
N	199

Note:

*p < .05, **p < .01, ***p < .001

Test 5

Logistic mixed effects model predicting betting in a yellow background session; the independent variables include potential reward value in the session, uncertainty level in the session, a dummy for treatment (reference: treatment C), and the previous decision (to control for choice inertia/stickiness). The main variables of interest are: - the reward variable; we expect a positive coefficient on that variable (Prediction 1). - the treatment variable; we expect a positive coefficient on this variable (Prediction 3). - the uncertainty variable; a positive coefficient would be consistent with Prediction 4.

Table 11: Logistic mixed effects model predicting betting in yellow background sessions for cravers (bet at least twice in yellow). Independent variables were reward value (0: low, 1: high), uncertainty (0: low, 1: high), treatment (0: control, 1: test), and previous choice (0: skip, 1: bet). Controls not shown were age, gender, major, and sequence number.

Variable	m_1
(Intercept)	-3.797 (0.301)***
reward_valueHigh	1.383 (0.07)***
uncertaintyHigh	0.681 (0.065)***
treatmentTest	0.264 (0.271)
sequence_number	-0.321 (0.032)***
previous_choice	1.316 (0.066)***
Nakagawa R^2	0.44
N	9224

Note:

*p < .05, **p < .01, ***p < .001

Test 5 with reward exposure and previous choice

Table 12: Test 5 with previous choice and reward history.

Variable	m_1
(Intercept)	-3.854 (0.306)***
reward_valueHigh	1.387 (0.07)***
uncertaintyHigh	0.683 (0.065)***
treatmentTest	0.374 (0.282)
sequence_number	-0.319 (0.032)***
previous_choice	1.312 (0.066)***
reward_exposure	-0.073 (0.051)
Nakagawa R^2	0.441
N	9224

Note:

*p < .05, **p < .01, ***p < .001

Next we show the VIF for the model with previous choice and reward exposure. VIF at 1 suggests there is no multicollinearity and VIF >10 suggests problematic collinearity.

Table 13: VIF for variables in test 5 with reward exposure included.

Variables	VIF
reward_value	1.027309
uncertainty	1.001925
treatment	2.678930
age	1.039347
gender	1.073227
major	1.024448
sequence_number	1.007021
previous_choice	1.049438
reward_exposure	2.625503

Stepwise model-building

This section shows the process of building the models from the theoretical version (in preregistration) to the model shown in the report. Models are built with incrementally more variables and each model is compared using Rˆ2.

Table 14: Stepwise comparison of models being built from theoretical base (1) to best fitting model (4).

Variable	(1)	(2)	(3)	(4)
(Intercept)	-3.485 (0.336)***	-3.552 (0.342)***	-3.797 (0.301)***	-3.854 (0.306)***
reward_valueHigh	1.552 (0.067)***	1.583 (0.068)***	1.383 (0.07)***	1.387 (0.07)***
uncertaintyHigh	0.744 (0.062)***	0.759 (0.063)***	0.681 (0.065)***	0.683 (0.065)***
treatmentTest	0.416 (0.305)	0.426 (0.31)	0.264 (0.271)	0.374 (0.282)
sequence_number	-	-0.349 (0.031)***	-0.321 (0.032)***	-0.319 (0.032)***
previous_choice	-	-	1.316 (0.066)***	1.312 (0.066)***
reward_exposure	-	-	-	-0.073 (0.051)
-	-	-	-	-
Rˆ2	0.427	0.447	0.44	0.441
N	9224	9224	9224	9224

Note:

*p < .05, **p < .01, **p < .001

Test 6

Paired t-test to compare participant betting rate in the high-uncertainty vs. low uncertainty sessions in the yellow background sessions, one-sided; H0: Betting rate is higher or equal in the low-uncertainty sessions. We expect to be able to reject H0 at a significance level of 5% (Prediction 4).

Paired t-test to compare participant betting rate in the high-uncertainty vs. low-uncertainty sessions in the yellow background sessions, one-sided; H0: Betting rate is higher or equal in the low-uncertainty sessions. We expect to be able to reject H0 at a 5% significance level.

Participant-level test ($N = 199$).

Data was skewed at 1.402 (ideal values within [-1, 1]) and a Shapiro-Wilks test showed non-normality. Data were Box-Cox transformed with lambda = -10. This reduced skew to 0.738 (acceptable value), but SW test still showed non-normality.

Paired t-test using Box-Cox transformed data showed higher betting rate in high uncertainty sessions ($t(198) = -3.542$, $p < .001$). Non-parametric paired Wilcoxon test using non-transformed data showed qualitatively same results ($V = 828.5$ $p < .001$).

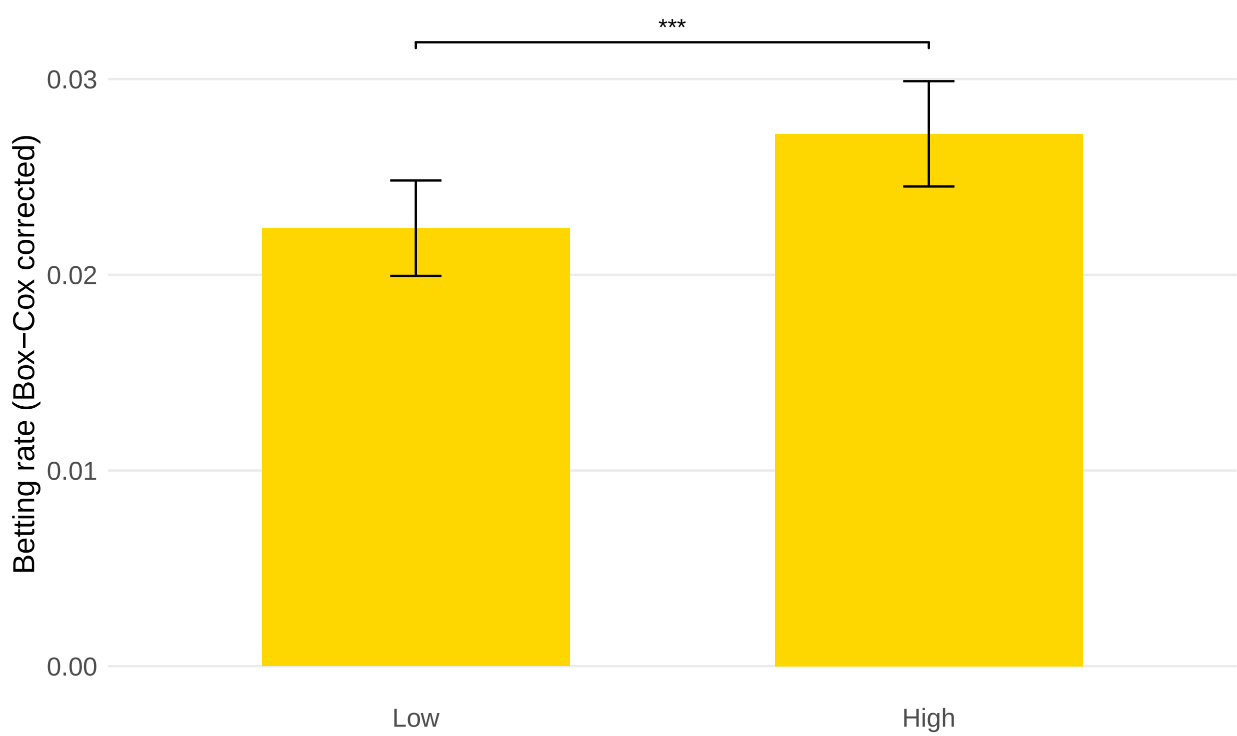

Figure 1: Average betting rate in yellow backround sessions for low and high uncertainty sessions for all participants ($t(198) = -3.542$, $p < .001$).

Test 7

T-test checking if betting rate in yellow session in the test treatment is different from 0. Participant-level test (N = 99). Additionally, shows comparison of betting rate in yellow between control and test treatment through two-sample t-test.

The t-test investigating whether betting rate in yellow in test was different from 0 was significant at, t(98) = 5.948, $p < .001$.

The t-test comparing betting in yellow between the control and test treatment was also significant, showing that betting rate was on average higher in test compared to control treatment (t(191.632) = -2.38, p = 0.018). This test was run on Box-Cox transformed data with lambda = -13.5.

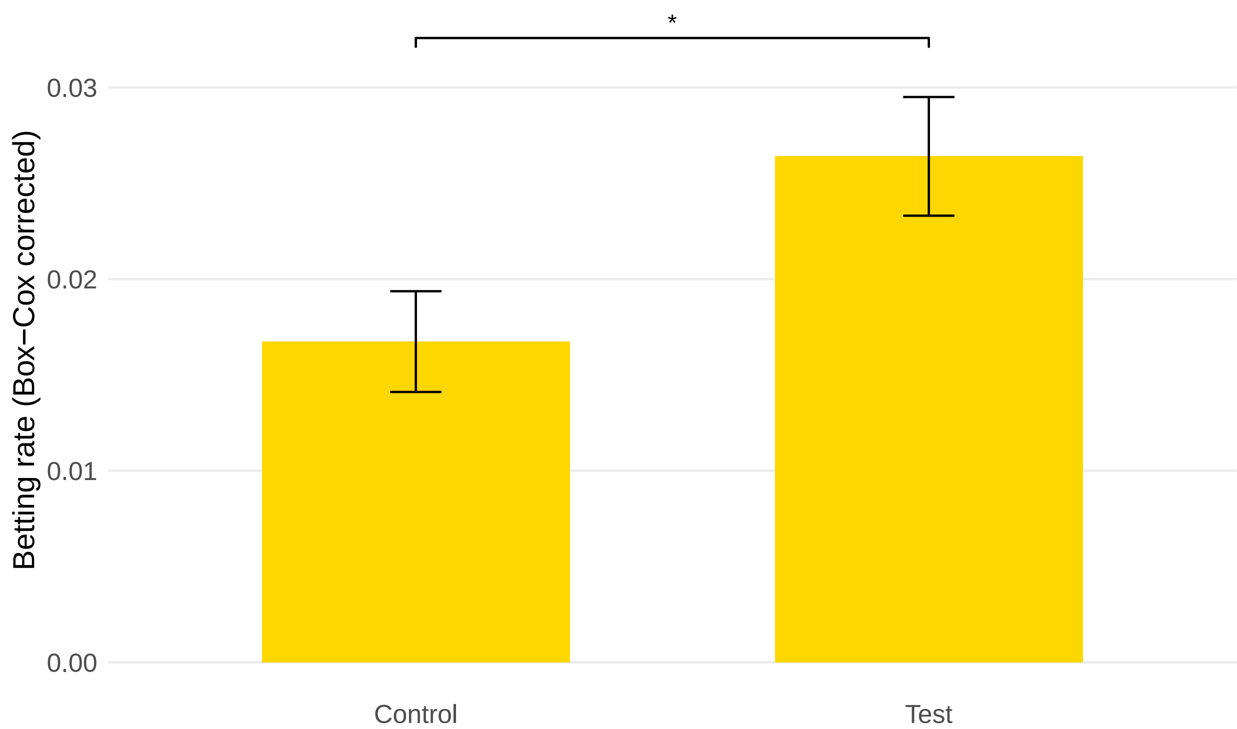

Figure 2: Average betting rate in yellow sessions for control and test treatment (t(191.632) = -2.38, p = 0.018).

Figures

S2 from OSF but with color as well - only for pooled

S4 but theta 1.0