Halfway Report 23-01-08

Samuel Thelaus

2023-01-08

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.

Four participants were excluded for not passing the requirement for MCQ questions, post-game quiz or questions after betting in yellow (I want to remember the odds guess questions are). Three excluded participants were in control treatment and one in the test treatment.

All excluded participants were cravers, as defined by betting at least twice in yellow.

Main analysis

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 113 participants after excluding participants.

Test 1

Paired one-tailed t-test checking if betting rate is equal or lower in high reward value trials. Participant-level test ($N = 113$).

Data was skewed at 1.69 (ideal values within [-1, 1]) and a Shapiro-Wilks test showed non-normality. Data were Box-Cox transformed with lambda = -19. This reduced skew to 0.909 (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 ($t(112) = -5.238$, $p < .001$). Non-parametric paired Wilcoxon test using non-transformed data showed qualitatively same results ($V = 84.5$ $p < .001$).

Test 2 & test 3

Logistic mixed-effects model predicting betting in all trials. Using penalized log-likelihood to correct for imbalanced groups. The same model is also presented with the previous choice variable to control for choice inertia. The models are compared on AIC (bias toward overfitting) and on BIC (bias toward underfitting). The forward stepwise process to build the model is shown in the appendix for these models and test 5.

```
## Warning in !is.null(rmarkdown::metadata$output) && rmarkdown::metadata$output
## %in% : 'length(x) = 2 > 1' in coercion to 'logical(1)'
```

Comparing the models, the difference between them was statistically significant ($Chi2(1) = 732.174$, $p < .001$) suggesting the model with previous choice is the better model.

Test 4

Test 5

Logistic mixed-effects model predicting betting among cravers in the yellow background sessions. Three models are shown. The first model uses all available data. The second model uses data only from the first yellow session in an order of yellow sessions for the control treatment. The third model uses data only from the first and second yellow session in an order of yellow sessions for the control treatment. The models can not be compared on AIC, BIC or with a chi-squared test as they are not using the same data.

Test 6

Bayesian t-test checking if betting rate in yellow session in the test treatment is different from 0. Participant-level test ($N = 57$).

Data was positively skewed at 1.016 (ideal values within [-1, 1]) and a Shapiro-Wilks test showed non-normality. Data were Box-Cox transformed with lambda = -4.75. This reduced skew to 0.565 (acceptable value), but SW test still showed non-normality.

Bayesian t-test on Box-Cox transformed betting rate showed a Bayes Factor of 4.398, with a probability of 81.5% that the mean was more than 0.

Table 1: Model 1 vs. model 2, with and without the previous choice variable

Variable	(1)	(2)
(Intercept)	-5.266 (0.385), p < .001***	-5.482 (0.346), p < .001***
reward_valueHigh	1.376 (0.058), p < .001***	1.236 (0.059), p < .001***
uncertaintyHigh	0.268 (0.051), p < .001***	0.22 (0.053), p < .001***
treatmentTest	1.027 (0.307), p = 0.001**	0.786 (0.276), p = 0.004**
colorBlue	10.805 (0.19), p < .001***	9.521 (0.189), p < .001***
age	-0.078 (0.155), p = 0.614	-0.061 (0.138), p = 0.66
gender2	-0.136 (0.306), p = 0.656	-0.123 (0.272), p = 0.651
gender3	-0.542 (1.424), p = 0.704	-0.462 (1.274), p = 0.717
major2	0.423 (0.333), p = 0.205	0.39 (0.296), p = 0.189
major3	0.145 (1.063), p = 0.892	0.157 (0.948), p = 0.869
major4	1.871 (1.073), p = 0.081	1.634 (0.952), p = 0.086
reaction_time	0.095 (0.026), p < .001***	0.025 (0.026), p = 0.337
sequence_number	-0.086 (0.026), p = 0.001**	-0.089 (0.027), p = 0.001**
treatmentTest:colorBlue	-3.091 (0.205), p < .001***	-2.716 (0.202), p < .001***
previous_choice	-	1.597 (0.057), p < .001***
-	-	-
AIC	12498	11767
BIC	12634	11913

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

Table 2: Three versions of test 5 on different subsets of data. Model 1 uses all available data. Model 2 uses only the first yellow session in an order in control. Model 3 uses only the first and second yellow session in an order in control.

Variable	(1)	(2)	(3)
(Intercept)	-3.562 (0.443), p < .001***	-3.54 (0.448), p < .001***	-3.566 (0.465), p < .001***
reward_valueHigh	1.503 (0.094), p < .001***	1.596 (0.108), p < .001***	1.548 (0.097), p < .001***
uncertaintyHigh	0.87 (0.088), p < .001***	0.771 (0.099), p < .001***	0.897 (0.091), p < .001***
treatmentTest	0.613 (0.371), p = 0.099	0.703 (0.379), p = 0.063	0.682 (0.391), p = 0.081
age	-0.311 (0.207), p = 0.133	-0.391 (0.214), p = 0.068	-0.326 (0.22), p = 0.138
gender2	-0.198 (0.361), p = 0.584	-0.296 (0.366), p = 0.419	-0.222 (0.381), p = 0.56
major2	-0.004 (0.449), p = 0.992	-0.128 (0.452), p = 0.776	-0.131 (0.47), p = 0.781
major4	1.348 (0.985), p = 0.171	1.229 (0.974), p = 0.207	1.207 (1.003), p = 0.229
reaction_time	0.485 (0.046), p < .001***	0.484 (0.052), p < .001***	0.463 (0.047), p < .001***
sequence_number	-0.147 (0.045), p = 0.001**	-0.166 (0.051), p = 0.001**	-0.142 (0.047), p = 0.002**
reward_history	-0.08 (0.064), p = 0.207	-0.051 (0.06), p = 0.398	-0.075 (0.063), p = 0.233

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

Test 7

Paired one-tailed t-test checking if betting rate is equal or lower in high uncertainty trials. Participant-level test.

Data was positively skewed at 0.818 (ideal values within [-1, 1]) and a Shapiro-Wilks test showed non-normality. Data were Box-Cox transformed with lambda = -8.75. This reduced skew to 0.818 (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 ($t(112) = -2.758$, $p = 0.003$). Non-parametric paired Wilcoxon test on non-transformed data showed qualitatively same results ($V = 232.5$ $p < .001$).

Plots

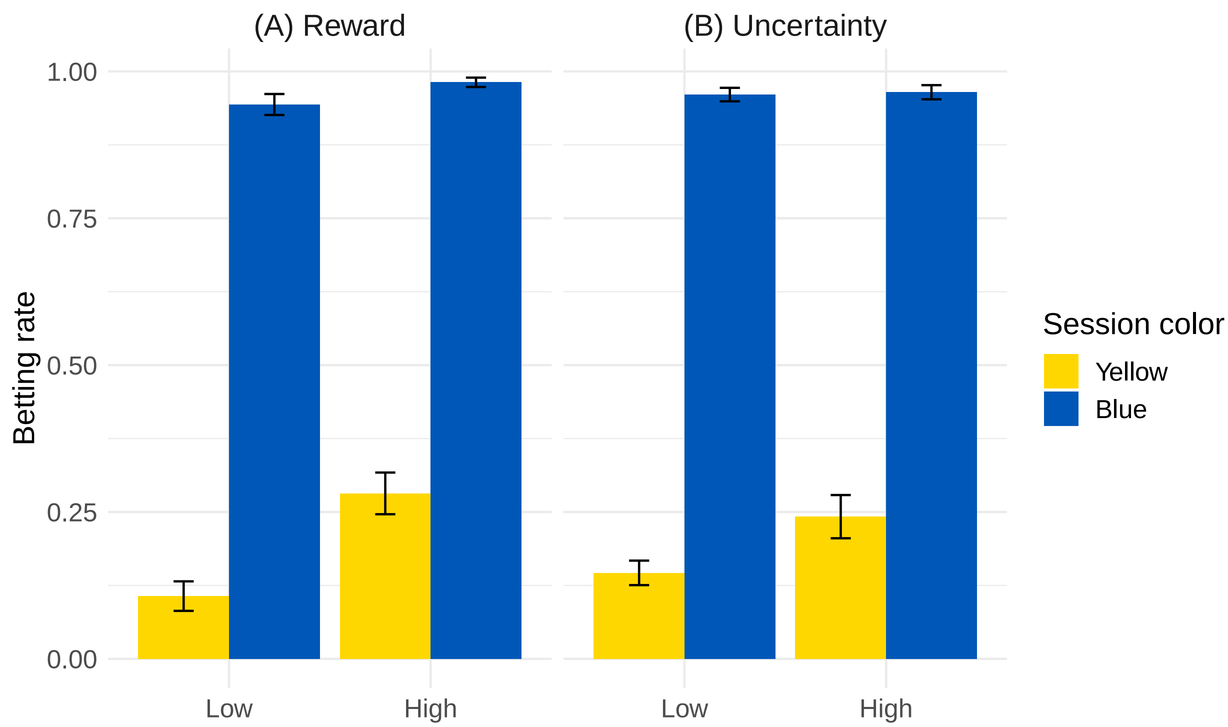

Figure 1: (A) Average betting rate in low and high reward sessions. (B) Average betting rate in low and high uncertainty sessions.

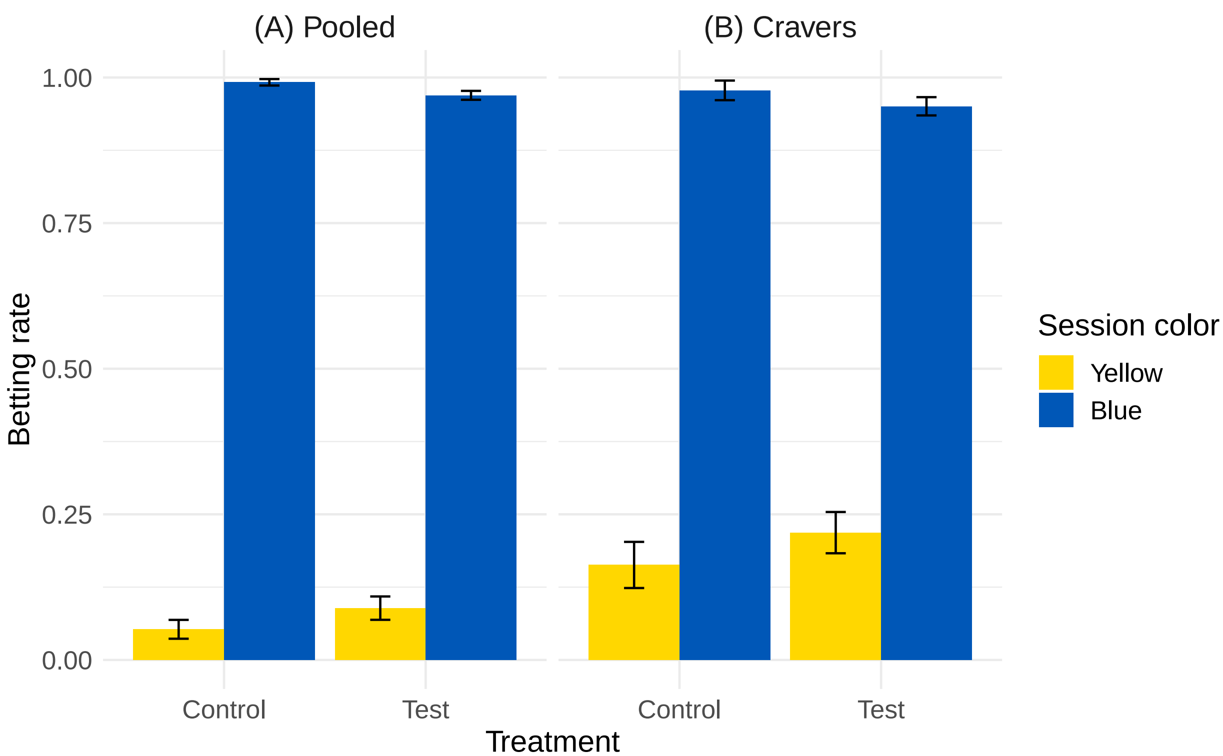

Figure 2: Average betting rate by session color and treatment, for all participants pooled and cravers individually.

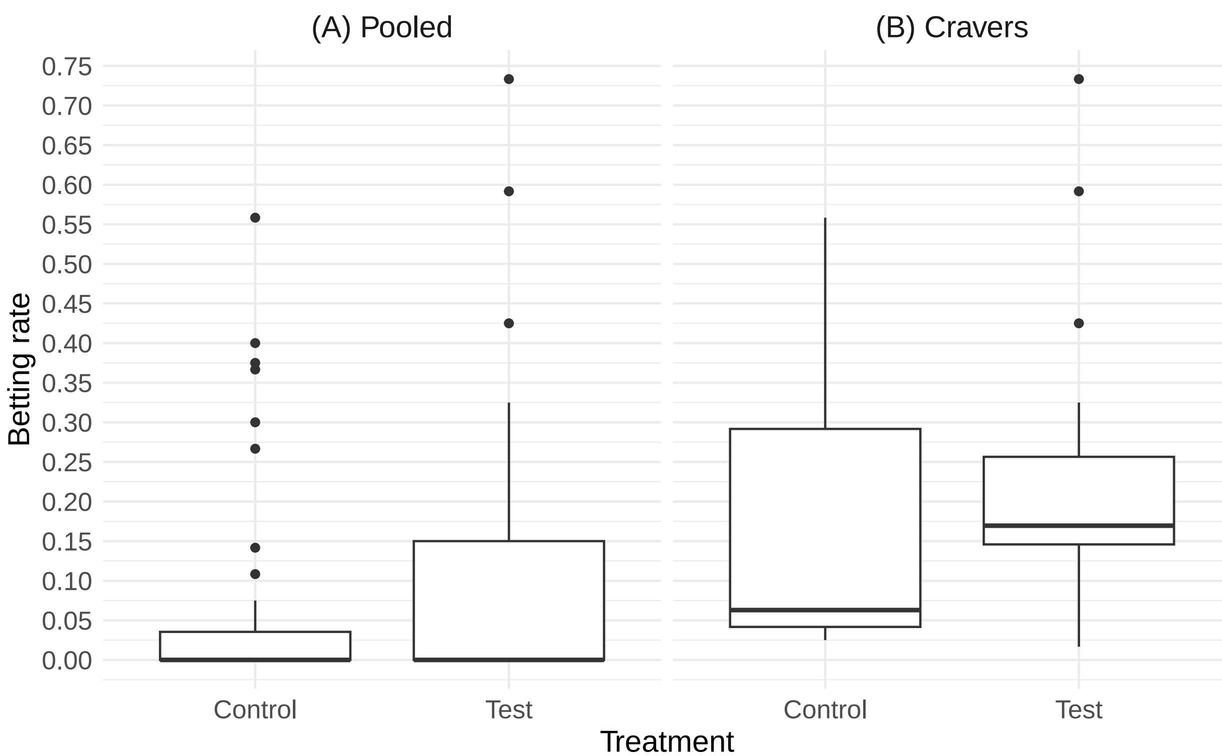

Figure 3: Distribution of betting rates by participant for all participants (A) and for the craver group (B) in yellow background sessions.

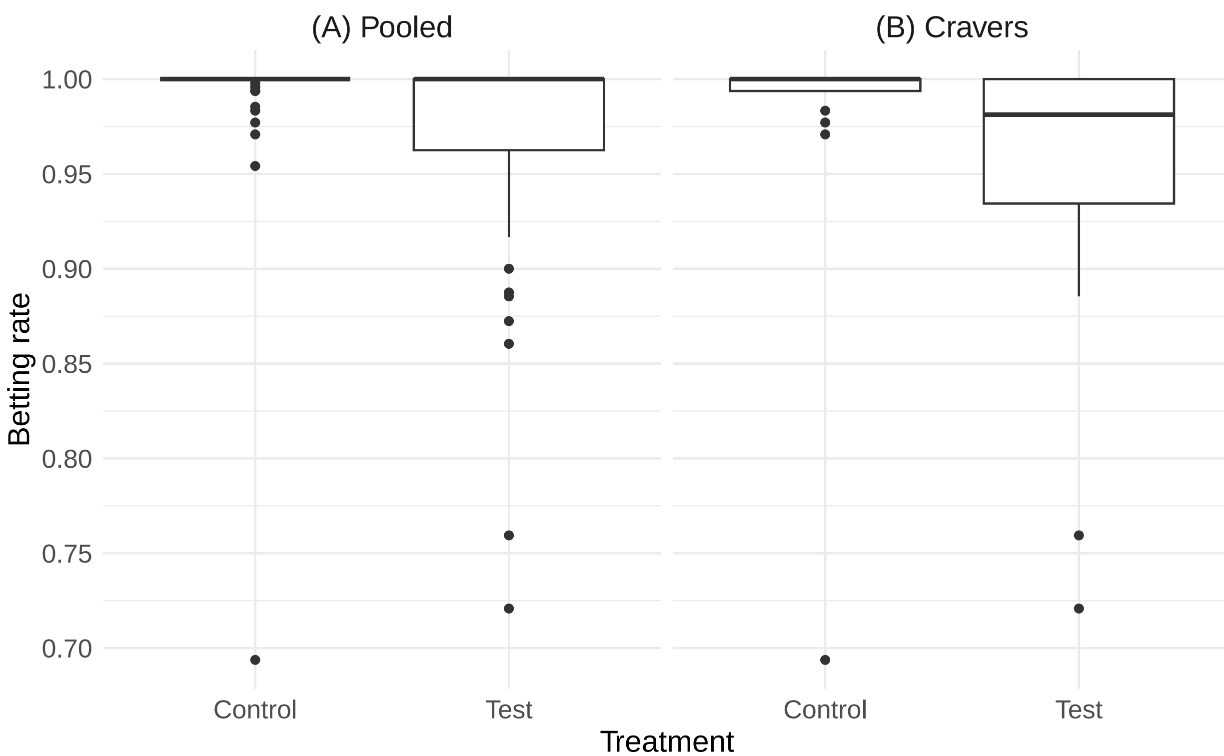

Figure 4: Distribution of betting rates by participant for all participants (A) and for the craver group (B) in blue background sessions.

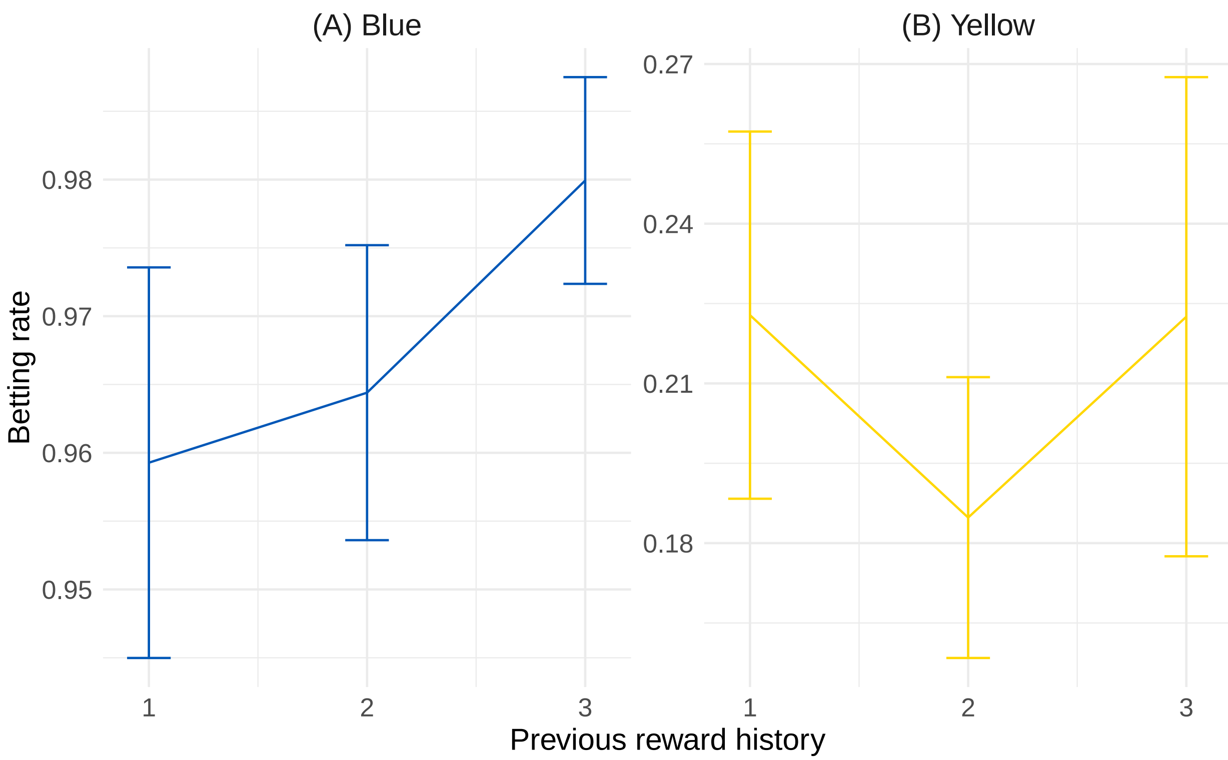

Figure 5: Betting rate in the blue background sessions (A) and yellow background sessions (B) as a function of prior reward history (as defined in computational model) for all participants and sessions, with reward history split into 3 equally sized bins.

Extra treatment/uncertainty checks

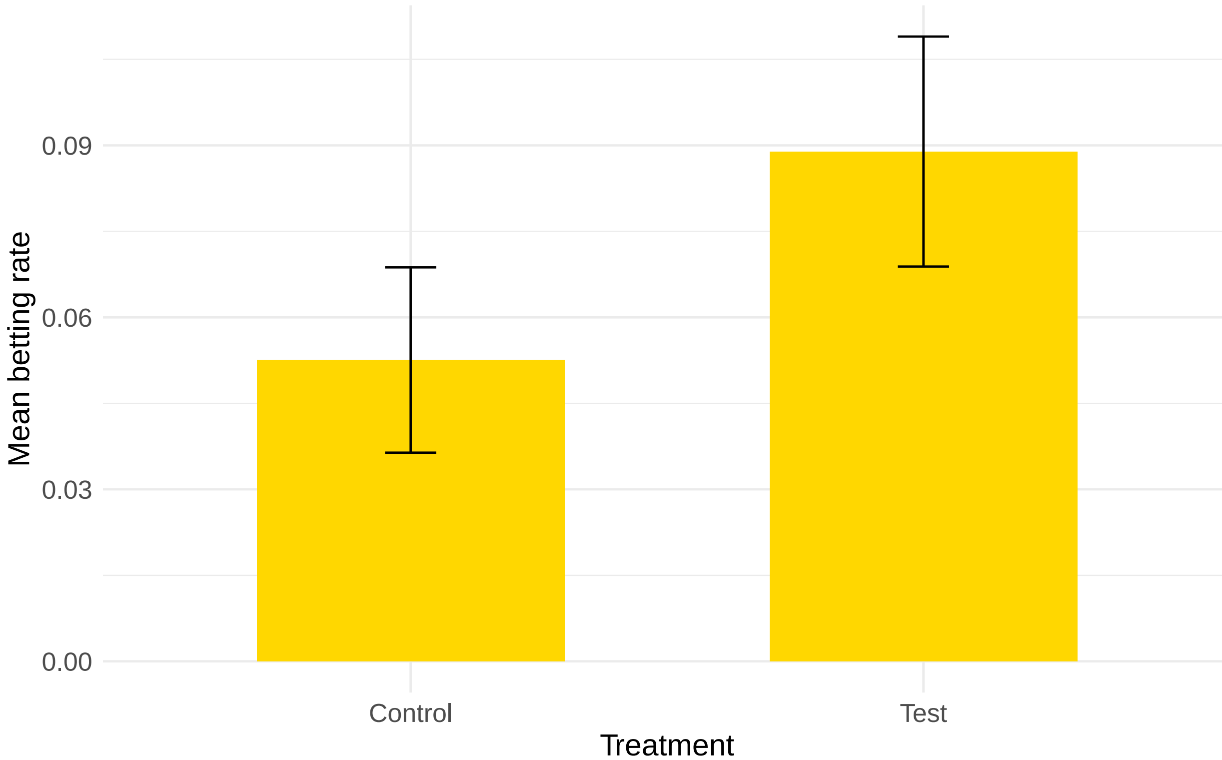

Figure 6: Average betting rate in yellow sessions for the control and test treatment.

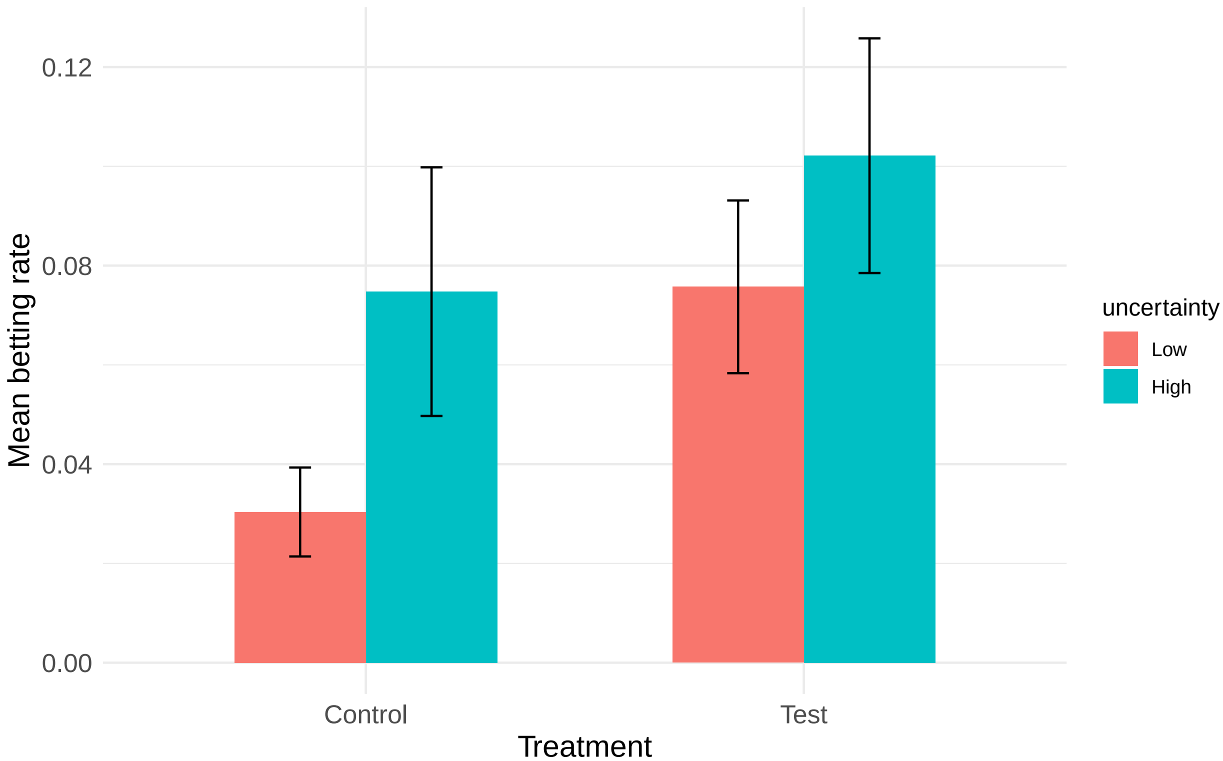

Figure 7: Average betting rates in control/test treatment for low/high uncertainty. There seems to be an effect of uncertainty in the control treatment but not in test.

Reaction time plots

Two plots showing reaction time by treatment and session color and in another plot by decision type and session color. Reaction time is also added as a variable in the logistic models in the main analysis above.

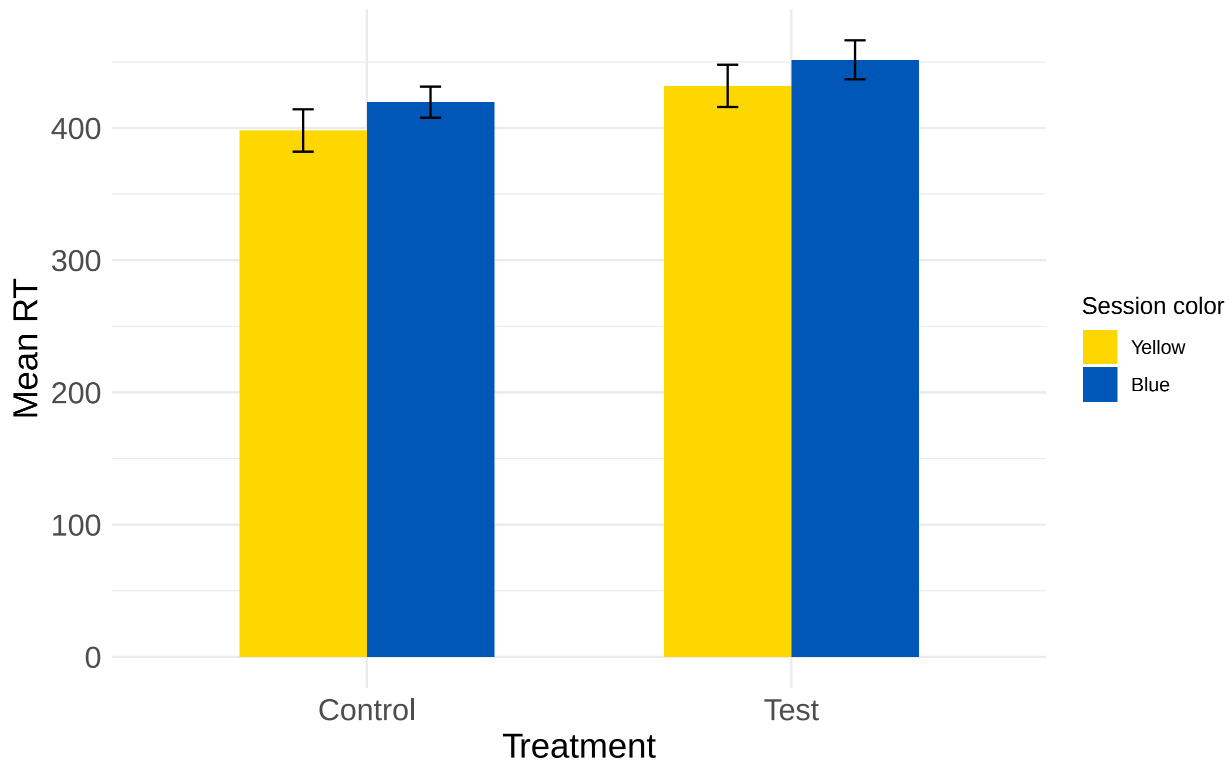

Figure 8: Average reaction time by session color and treatment.

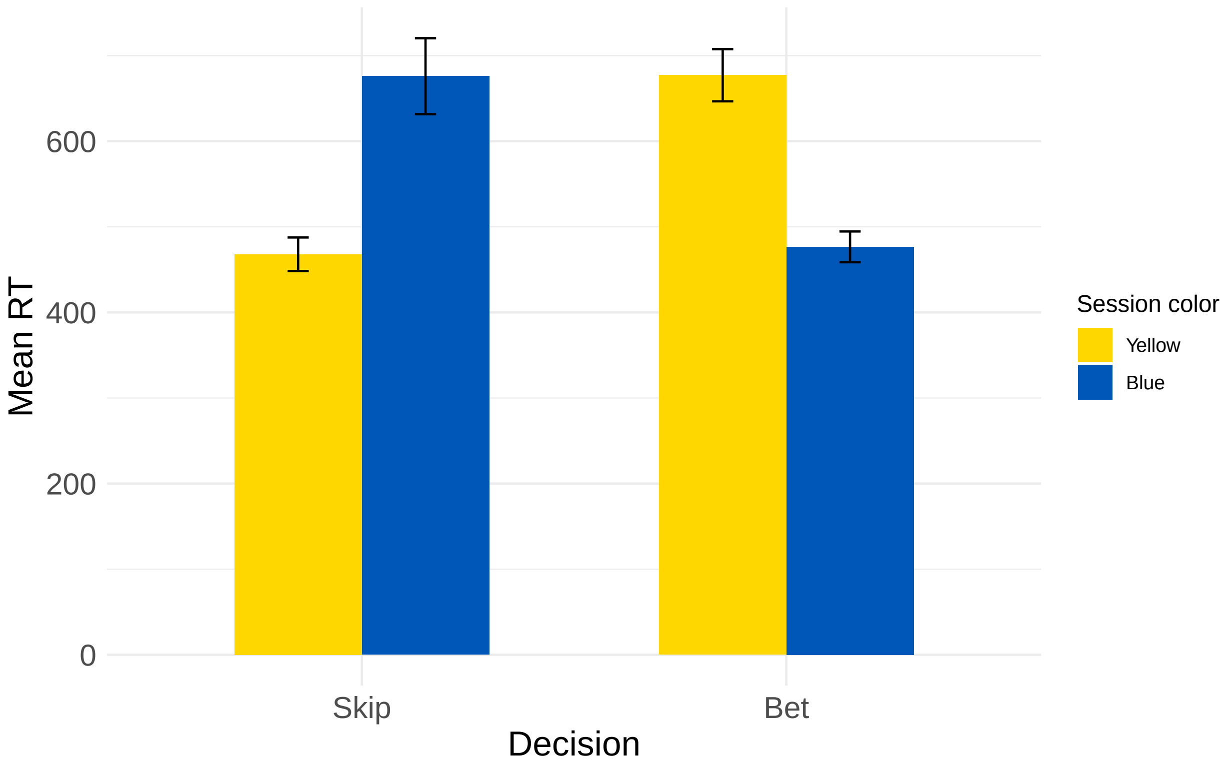

Figure 9: Average RT by decision (skip/bet) and session color (blue/yellow).

Reward history with different theta

This section checks results for the reward history variable when using theta = 0.95 and theta = 1 instead of 0.9. This is done by reproducing test 5 as well as figure 5 using the new design of the reward history variable. Test 5 and the versions with theta = 0.95 and 1 are shown in the same table.

Test 5 copy

Table 3: Three versions of test 5 with different values of theta. Model 1 is the original version with theta = 0.90. Model 2 has theta 0.95 and model 3 has theta 1.00.

Variable	(1)	(2)	(3)
(Intercept)	-3.562 (0.443), p < .001***	-3.562 (0.443), p < .001***	-3.562 (0.443), p < .001***
reward_valueHigh	1.503 (0.094), p < .001***	1.503 (0.094), p < .001***	1.503 (0.094), p < .001***
uncertaintyHigh	0.87 (0.088), p < .001***	0.87 (0.088), p < .001***	0.87 (0.088), p < .001***
treatmentTest	0.613 (0.371), p = 0.099	0.613 (0.371), p = 0.099	0.613 (0.371), p = 0.099
age	-0.311 (0.207), p = 0.133	-0.311 (0.207), p = 0.133	-0.311 (0.207), p = 0.133
gender2	-0.198 (0.361), p = 0.584	-0.198 (0.361), p = 0.584	-0.198 (0.361), p = 0.584
major2	-0.004 (0.449), p = 0.992	-0.004 (0.449), p = 0.992	-0.004 (0.449), p = 0.992
major4	1.348 (0.985), p = 0.171	1.348 (0.985), p = 0.171	1.348 (0.985), p = 0.171
reaction_time	0.485 (0.046), p < .001***	0.485 (0.046), p < .001***	0.485 (0.046), p < .001***
sequence_number	-0.147 (0.045), p = 0.001**	-0.147 (0.045), p = 0.001**	-0.147 (0.045), p = 0.001**
reward_history	-0.08 (0.064), p = 0.207	-0.08 (0.064), p = 0.207	-0.08 (0.064), p = 0.207

Note:

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

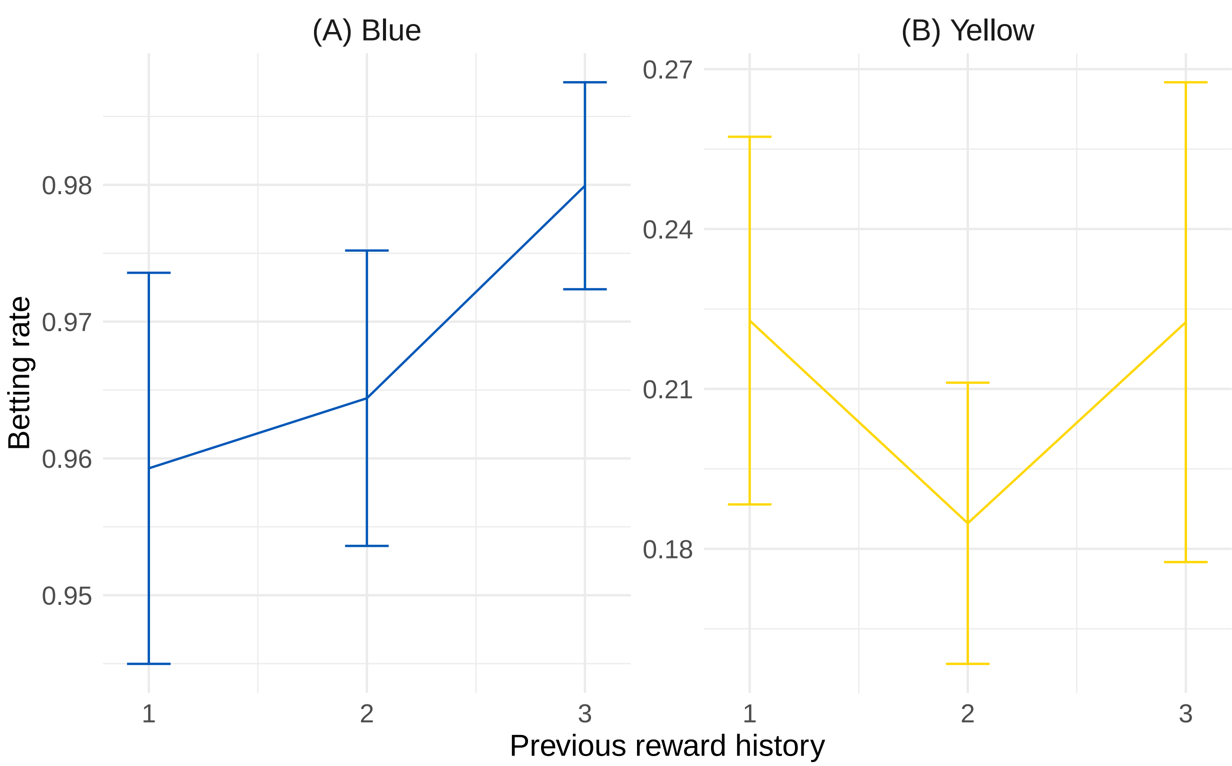

Figure 10: Betting rate in the blue background sessions (A) and yellow background sessions (B) as a function of prior reward history (as defined in computational model) for cravers in all sessions, with reward history split into 3 equally sized bins. In this plot, theta is 0.95.

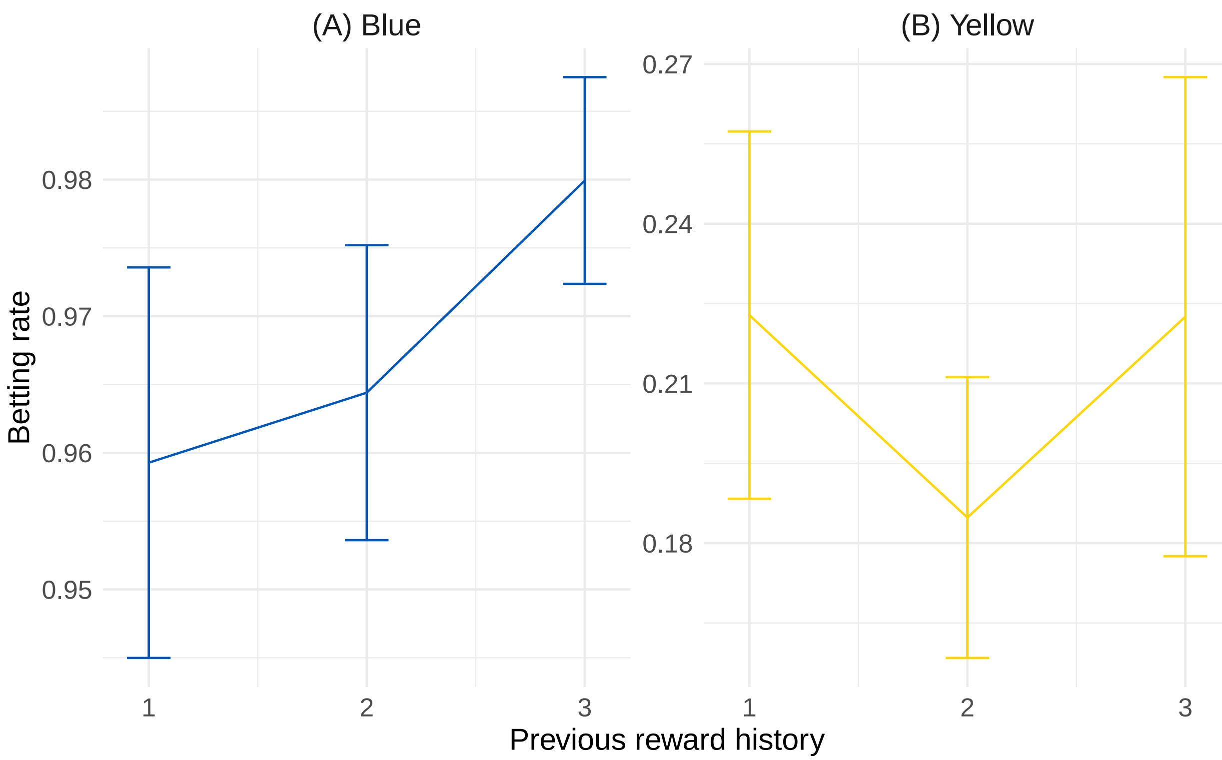

Figure 11: Betting rate in the blue background sessions (A) and yellow background sessions (B) as a function of prior reward history (as defined in computational model) for cravers in all sessions, with reward history split into 3 equally sized bins. In this plot, theta is 1.00.

Controlling for losses

We wanted to see if there was a correlation between the losses people accrued and the betting rate in yellow background sessions. This was meant to represent whether people learned from their mistakes. This was done in a few steps.

First, we checked overall betting rates in yellow over the 6 sequences for control and test. This was simply to see betting over the course of the experiment in general, as a baseline.

Next, we checked two different variables, both related to losses. First, we checked betting rate as a function of cumulative losses (in yellow only). Then, we checked betting rate as a function of cumulative losses in yellow but dividing by the cumulative bets in yellow. In this way, a participant who bets a lot in yellow but luckily does not lose a lot, will have a lower "loss score".

Finally, we plot betting rates in yellow each sequence for participants who with low, medium and high losses in the first sequence. This is meant to show whether high losses in the beginning of the experiment result in different betting patterns throughout.

These four checks are presented in the figures below. We importantly only used cravers for this analysis as others did not bet in yellow more than once.

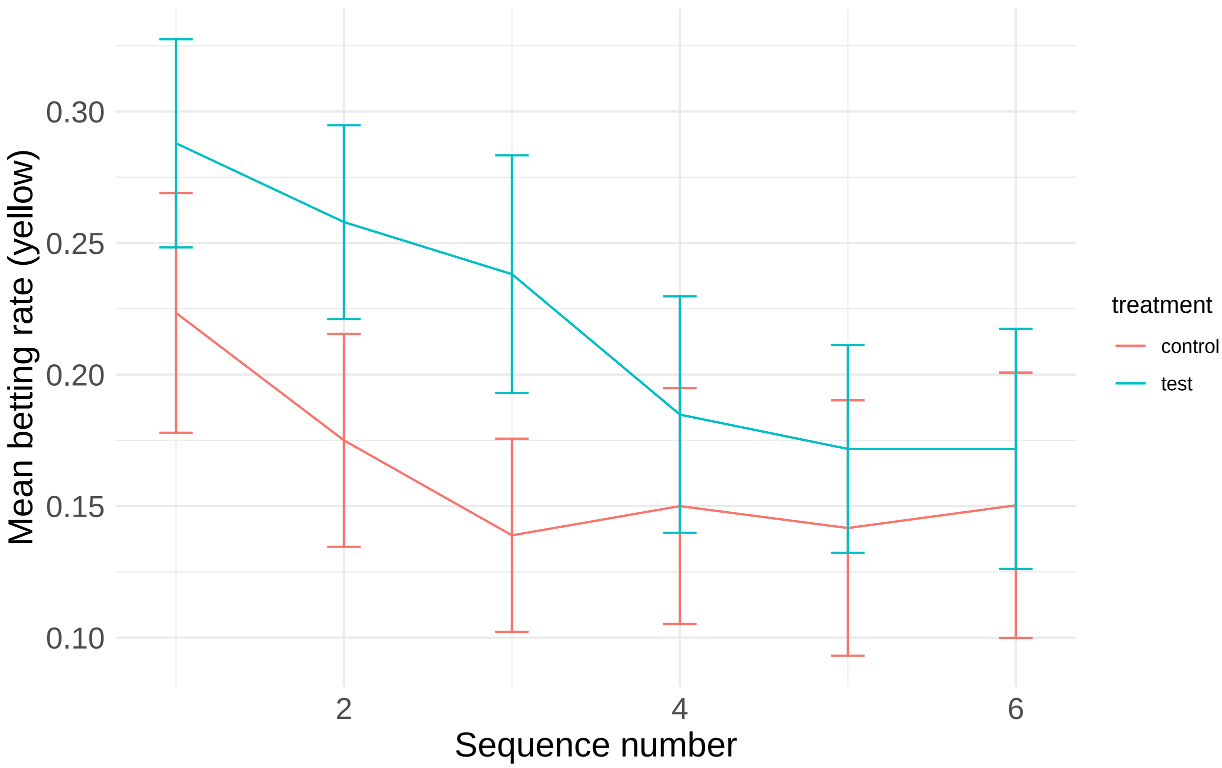

Figure 12: Betting rate in yellow as a function of sequence number in test and control. Error bars are SEM.

Next, we include a new variable in test 5. We introduce cumulative losses to the model and remove the reward history variable (due to colinearity).

To see if the effect differs by treatment, we also include the interaction between treatment and the loss variable.

The interaction is negative, meaning that the effect of accrued losses on betting is lower in the test treatment compared to the control treatment. In other words, participants overall bet more if they have lost more, but this is more pronounced in the control treatment.

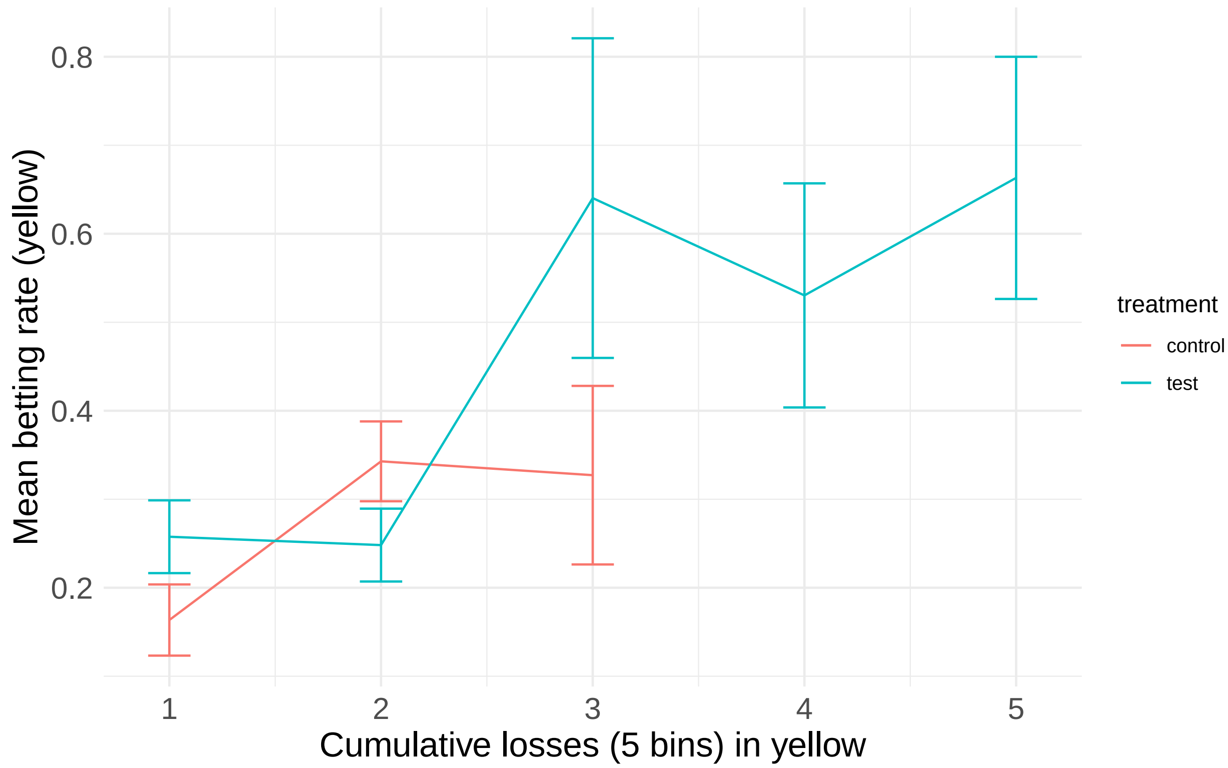

Figure 13: Betting rate in yellow by cumulative losses (in 5 bins) for test and control. Error bars are SEM.

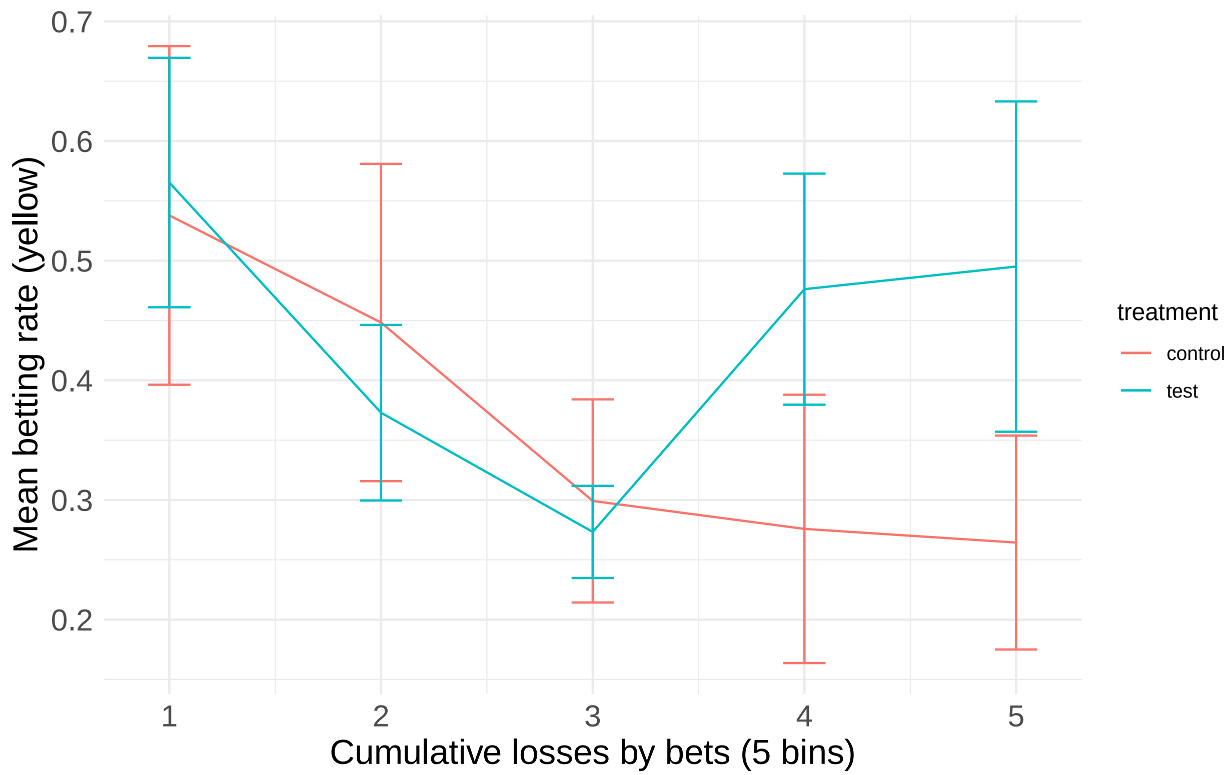

Figure 14: Betting rate by 5 bins representing cumulative losses over cumulative bets. Error bars are SEM.

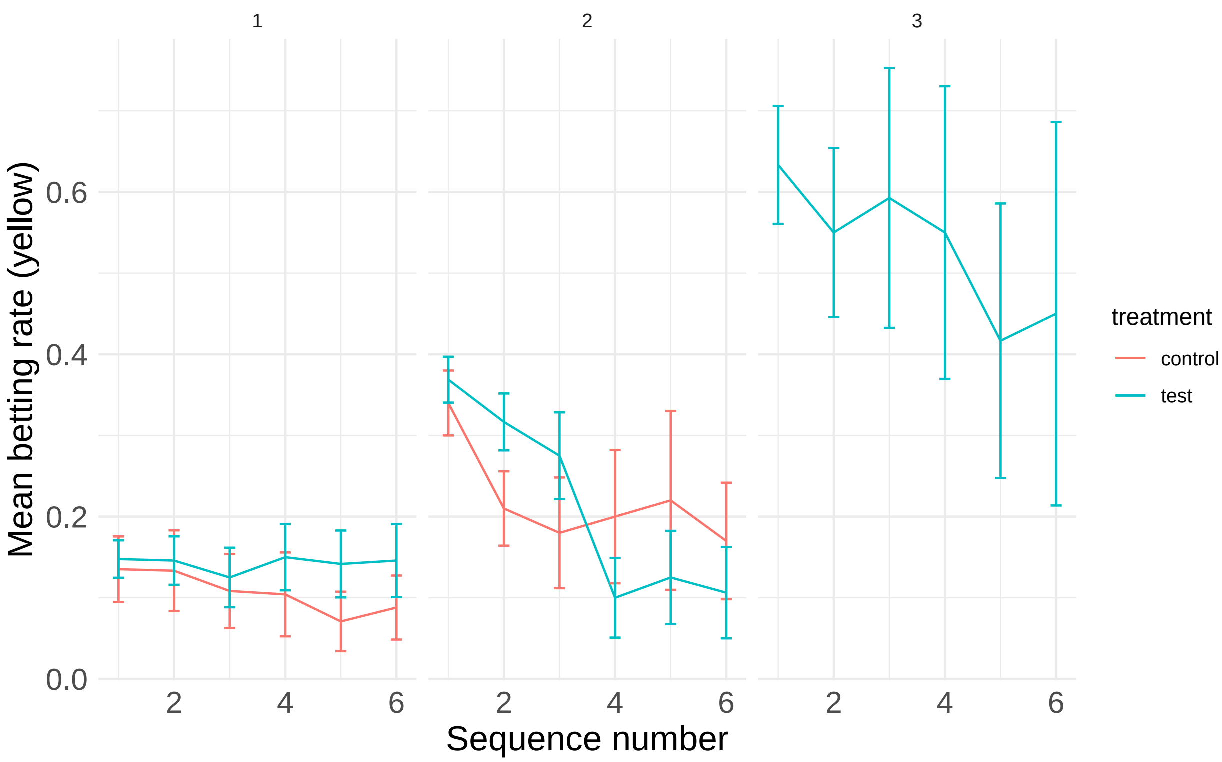

Figure 15: Betting rate in yellow by sequence number. The plot is separated in 3 increasing panes with low (1), medium (2), and high (3) losses in the first sequence.

Test 5 - first and second in sequence only

Variable	m_1
(Intercept)	-3.355 (0.401), p < .001***
reward_valueHigh	1.49 (0.094), p < .001***
uncertaintyHigh	0.867 (0.088), p < .001***
treatmentTest	0.254 (0.329), p = 0.44
age	-0.282 (0.187), p = 0.131
gender2	-0.218 (0.326), p = 0.503
major2	0.029 (0.408), p = 0.944
major4	1.184 (0.909), p = 0.193
losses	0.941 (0.184), p < .001***
reaction_time	0.489 (0.046), p < .001***
sequence_number	-0.933 (0.176), p < .001***
treatmentTest:losses	-0.229 (0.087), p = 0.009**

Note:

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

First 4 sessions of blue in test/control

We checked the difference in betting rates in the first 4 blue sessions in the test and control treatment separately. This is shown in a bar chart below as well as a t-test checking the difference in betting rates between test and control. Only the first “actual” blue sessions are used. Interspersed sessions and clocks before the experiment (to avoid boredom) were not used.

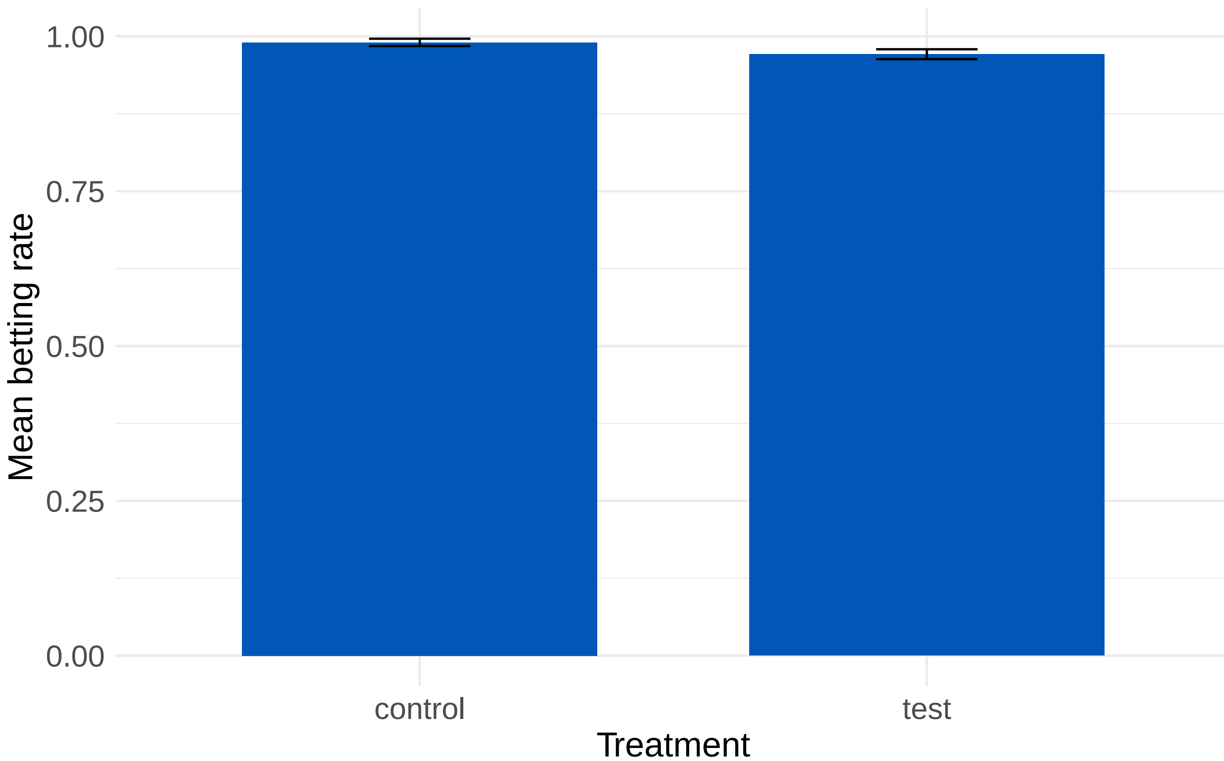

Figure 16: Average betting rates in the first 4 blue sessions for the control and test treatment. Error bars show SEM.

A two-sample t-test comparing the difference in betting rates in the first 4 blue sessions found no difference in mean between the two treatments (Welch two-sample t-test, $t(111) = 1.912$, $p = 0.058$).

Table 4: Three versions of test 5 on different subsets of data. Model 1 uses all available data. Model 2 uses only the first yellow session in an order in control. Model 3 uses only the first and second yellow session in an order in control.

Variable	(1)	(2)	(3)
(Intercept)	-3.241 (0.542), p < .001***	-3.352 (0.572), p < .001***	-3.339 (0.585), p < .001***
reward_valueHigh	1.521 (0.1), p < .001***	1.64 (0.118), p < .001***	1.57 (0.104), p < .001***
uncertaintyHigh	0.882 (0.095), p < .001***	0.756 (0.108), p < .001***	0.914 (0.099), p < .001***
treatmentTest	0.645 (0.421), p = 0.125	0.754 (0.445), p = 0.09	0.775 (0.456), p = 0.089
age	-0.137 (0.22), p = 0.532	-0.194 (0.231), p = 0.4	-0.105 (0.238), p = 0.66
gender2	-0.16 (0.41), p = 0.697	-0.272 (0.435), p = 0.532	-0.239 (0.445), p = 0.591
major2	-0.321 (0.51), p = 0.528	-0.329 (0.533), p = 0.538	-0.37 (0.547), p = 0.499
major4	1.198 (1.018), p = 0.239	1.098 (1.037), p = 0.289	1.122 (1.057), p = 0.288
reaction_time	0.436 (0.05), p < .001***	0.433 (0.057), p < .001***	0.412 (0.052), p < .001***
sequence_number	-0.218 (0.049), p < .001***	-0.269 (0.056), p < .001***	-0.222 (0.051), p < .001***
reward_history	-0.11 (0.07), p = 0.117	-0.062 (0.067), p = 0.353	-0.097 (0.07), p = 0.167

Note:

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

Appendix

Treatment effect in new sample only

We also test the treatment effect in yellow in the new sample when we exclude all pilot participants. This subsample has 95 participants.

The treatment effect is examined by running test 5 and producing figure 2 for the new sample alone.

Test 5

Logistic mixed-effects model predicting betting among cravers in the yellow background sessions. Three models are shown. The first model uses all available data. The second model uses data only from the first yellow session in an order of yellow sessions for the control treatment. The third model uses data only from the first and second yellow session in an order of yellow sessions for the control treatment. The models can not be compared on AIC, BIC or with a chi-squared test as they are not using the same data.

Boredom in control in yellow

We wanted to check whether participants might bet in yellow in the control treatment out of boredom (skipping many times in a row could be boring as nothing happens). To do this, we fashioned a few extra variables and checked their importance in explaining differences in betting rates.

First, we produced two additional plots to check whether betting was different over the course of sessions in yellow in control. One variable compared the first session in a sequence of yellow to the second and third, and one variable compared the first and second session to the third in a sequence. The results are shown in the bar charts below.

Stepwise model building process

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

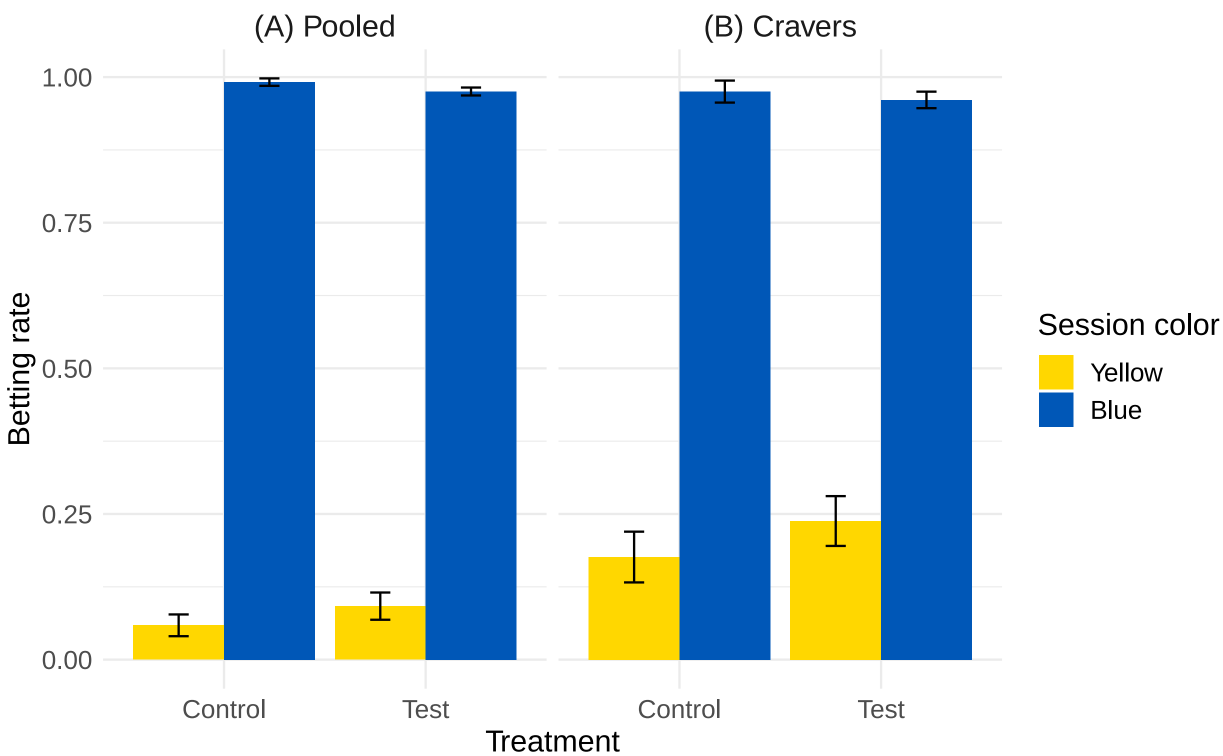

Figure 17: Bar chart showing betting rate in control and test treatment in yellow and blue background sessions. The panes represent data for all participants (A) and only cravers (B).

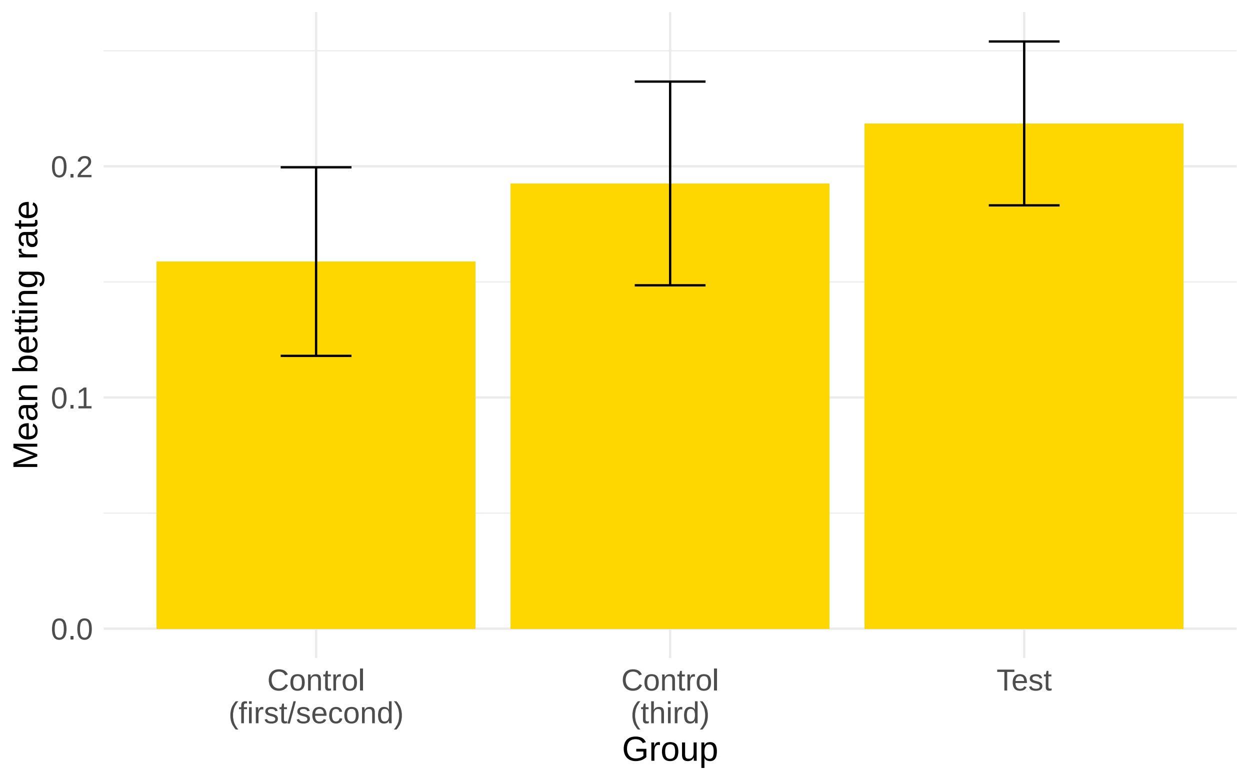

Figure 18: Betting rate in the first/second or third sessions in a sequence of yellow sessions in the control treatment. Betting rate for the test treatment is shown individually.

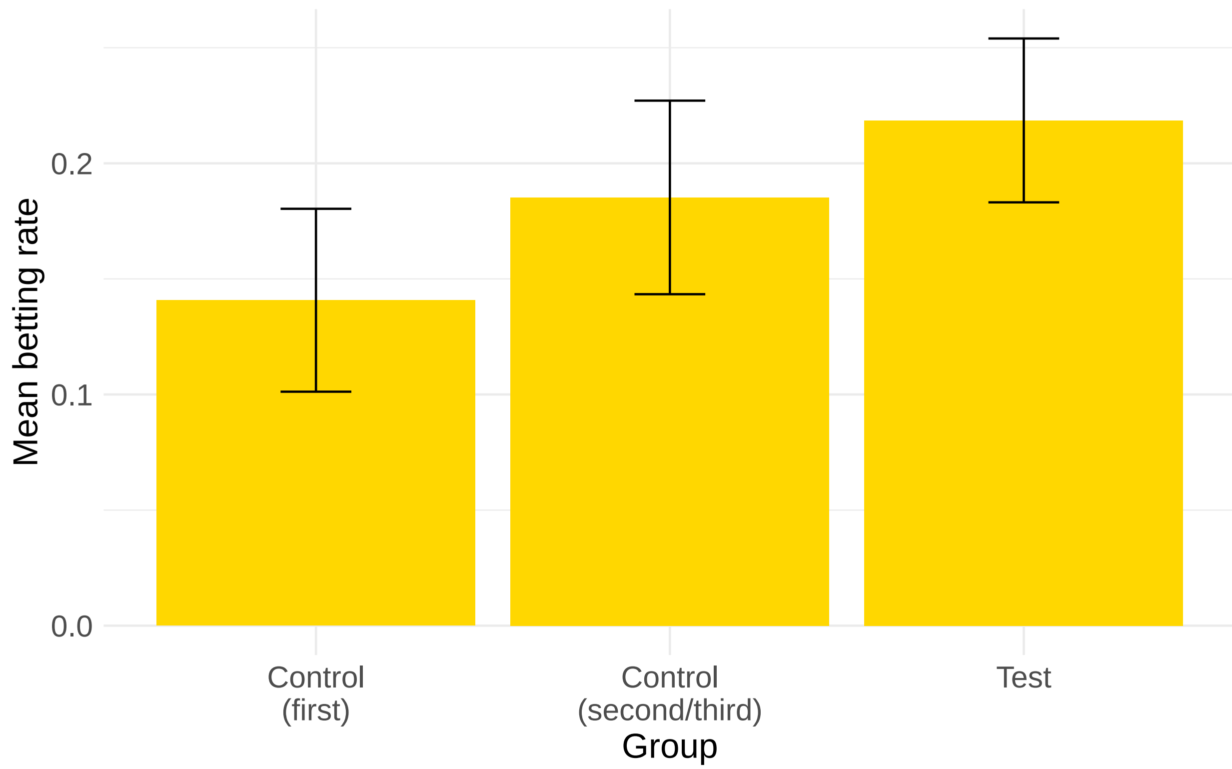

Figure 19: Betting rate in the first or second/third session in a sequence of yellow sessions in the control treatment. Betting rate for the test treatment is shown individually.

Table 5: 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)	
(Intercept)	-5.253 (0.393), p < .001***	-5.207 (0.383), p < .001***	-5.266 (0.385), p < .001***	-5.48
reward_valueHigh	1.374 (0.057), p < .001***	1.37 (0.057), p < .001***	1.376 (0.058), p < .001***	1.236
uncertaintyHigh	0.259 (0.051), p < .001***	0.262 (0.051), p < .001***	0.268 (0.051), p < .001***	0.22
treatmentTest	1.041 (0.311), p = 0.001**	1.002 (0.306), p = 0.001**	1.027 (0.307), p = 0.001**	0.786
colorBlue	10.811 (0.19), p < .001***	10.786 (0.189), p < .001***	10.805 (0.19), p < .001***	9.521
age	-0.076 (0.156), p = 0.625	-0.086 (0.155), p = 0.579	-0.078 (0.155), p = 0.614	-0.0
gender2	-0.174 (0.31), p = 0.574	-0.148 (0.305), p = 0.626	-0.136 (0.306), p = 0.656	-0.1
gender3	-0.661 (1.448), p = 0.648	-0.547 (1.428), p = 0.702	-0.542 (1.424), p = 0.704	-0.4
major2	0.438 (0.337), p = 0.194	0.386 (0.332), p = 0.244	0.423 (0.333), p = 0.205	0.3
major3	0.176 (1.078), p = 0.871	0.088 (1.063), p = 0.934	0.145 (1.063), p = 0.892	0.15
major4	1.906 (1.088), p = 0.08	1.838 (1.074), p = 0.087	1.871 (1.073), p = 0.081	1.63
treatmentTest:colorBlue	-3.107 (0.206), p < .001***	-3.082 (0.205), p < .001***	-3.091 (0.205), p < .001***	-2.71
reaction_time	-	0.113 (0.025), p < .001***	0.095 (0.026), p < .001***	0.02
sequence_number	-	-	-0.086 (0.026), p = 0.001**	-0.08
previous_choice	-	-	-	1.597
-	-	-	-	
AIC	12525	12506	12498	
BIC	12643	12634	12634	

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

using AIC and BIC.

Test 3 model (includes test 2)

Test 5 model

Table 6: Stepwise comparison of models being built from theoretical base (1) to best fitting model (4).

Variable	(1)	(2)	(3)	
(Intercept)	-3.451 (0.455), p < .001***	-3.497 (0.436), p < .001***	-3.515 (0.439), p < .001***	-3.562 (0.44
reward_valueHigh	1.504 (0.091), p < .001***	1.494 (0.093), p < .001***	1.506 (0.094), p < .001***	1.503 (0.09
uncertaintyHigh	0.822 (0.085), p < .001***	0.881 (0.088), p < .001***	0.881 (0.088), p < .001***	0.87 (0.088
treatmentTest	0.558 (0.372), p = 0.133	0.483 (0.355), p = 0.174	0.497 (0.359), p = 0.166	0.613 (0.3
age	-0.29 (0.214), p = 0.174	-0.299 (0.204), p = 0.143	-0.304 (0.206), p = 0.139	-0.311 (0.2
gender2	-0.182 (0.372), p = 0.624	-0.185 (0.356), p = 0.604	-0.185 (0.359), p = 0.606	-0.198 (0.3
major2	0.036 (0.467), p = 0.939	0.005 (0.446), p = 0.992	0 (0.448), p = 1	-0.004 (0.4
major4	1.553 (1.067), p = 0.146	1.436 (1.01), p = 0.155	1.392 (0.999), p = 0.163	1.348 (0.9
reaction_time	-	0.526 (0.044), p < .001***	0.478 (0.046), p < .001***	0.485 (0.04
sequence_number	-	-	-0.151 (0.045), p = 0.001**	-0.147 (0.04
reward_history	-	-	-	-0.08 (0.0
-	-	-	-	
AIC	3761	3613	3604	
BIC	3820	3678	3675	

Note:

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