Supplementary Material

Social threat avoidance depends on action-outcome predictability

Matteo Sequestro[1*], Jade Serfaty[1], and Julie Grèzes[1#*], Rocco Mennella[1,2#]

[1] Cognitive and Computational Neuroscience Laboratory (LNC[2]), Inserm U960, Department of Cognitive Studies, École Normale Supérieure, PSL University, Paris, France.

[2] Laboratoire des Interactions Cognition Action Émotion (LICAÉ, EA2931), Université Paris Nanterre (UFR STAPS), Nanterre, France.

[#]Julie Grèzes and Rocco Mennella should be considered joint senior authors.

[*]Corresponding authors

Matteo Sequestro (matteo.sequestro@gmail.com)

Julie Grèzes (julie.grezes@ens.psl.eu)

A. Pilot Experiment

Materials and Methods

Participants

Thirty-three volunteers participated in this pilot. Three participants have been excluded from the analyses because they had to interrupt the experiment due cybersickness symptoms, as well as 7 participants who did not interrupt the experiment but still have reported high cybersickness symptoms. This left us with a sample of 23 participants (15 females; mean age ± SD = 24.91 ± 5.07). Participants were left- (17%) or right-handed (83%), and they had no history of neurological or psychiatric disorder (see *Supplementary Material B* for a full list of inclusion criteria). The experimental protocol was approved by INSERM, licensed by the local research ethics committee (IRB00003888 – Avis 18-544-ter - 25.10.2021) and carried out in accordance with the Declaration of Helsinki. Participants provided informed written consent and were compensated for their participation.

Stimuli

The stimuli were identical to experiments 1, 2 and 3.

Task

The task design used for this pilot was almost identical to the one described in the *Experiment 1* section, except for two differences: (a) Early choices (i.e. choices before the doors' opening) were not followed by the text prompt saying: "*Ne pas choisir avant l'ouverture des portes!*" ("*Do not choose before the doors opening!*") and the trial was not restarted; (b) trials ended with participants' view rotating toward the elevators' door to face the avatar. This rotation was removed from the final version of the task as it caused cybersickness symptoms.

Equipment

The VR HMD and controllers were the same as in the other experiments. Differently from main experiments, the software of this pilot ran on a Razer Blade 17 laptop with 2.3GHz 8-core Intel i7-11800H processor, a 32GB DDR4 3200 MHz RAM, and a NVIDIA GeForce RTX 3070 8GB GDDR6 graphic card.

Data Analysis

The analyses were identical to experiments 1 to 3. Eleven trials were excluded due to a response time lower than 150 ms (0.80%).

Results

The best fitting model has been found to be the one including the Condition, the RTs (grand average centered) and their interaction as predictors. Participants avoided the angry display more in the Predictable condition than in the Unpredictable one (Condition effect: β=.429; 95%CrI=[.141, .707]; $p_{<0}$=.002; BF_{10}=4.401) while avoidance probability was at a chance level in the Unpredictable condition (Intercept: 95%CrI=[-.054, .479]; BF_{01}=5.716). Both the main effect of RTs and their interaction effect with Condition were found to be credible (RTs effect: β=.728; 95%CrI=[.011, 1.452]; $p_{<0}$=.024; BF_{10}=1.087; Condition x RTs interaction: β=1.108; 95%CrI=[.078, 2.107]; $p_{<0}$=.017; BF_{10}=1.897).

Table A1: *Pilot experiment regression table*

	Estimate	Est.Error	95%CrI-low	95%CrI-Up	$p_{>0}$	$p_{<0}$	BF_{01}	BF_{10}	R-hat	ESS
Inter	.208	.135	-.054	.479	.939	.061	5.716	.175	1.001	10593
Cond	.429	.141	.148	.707	.998	.002	.227	4.401	1	18733
RTs	.728	.366	.011	.452	.976	.024	.920	1.087	1	17141
Cond:RTs	1.108	.518	.078	2.107	.983	.017	.527	1.897	1	16738

Note: Cond: Condition effect. Est.Error: Estimate Error. 95%CrI-low and -up: lower and upper boundary of the 95% Credible Interval. $p_{>0}$ and $p_{<0}$: Proportion of posterior samples greater and lower than 0. BF_{01} and BF_{10}: Bayes Factor for the parameter to be equal or different from 0. Rhat: R-hat potential scale reduction factor. ESS: Effective Sample Size.

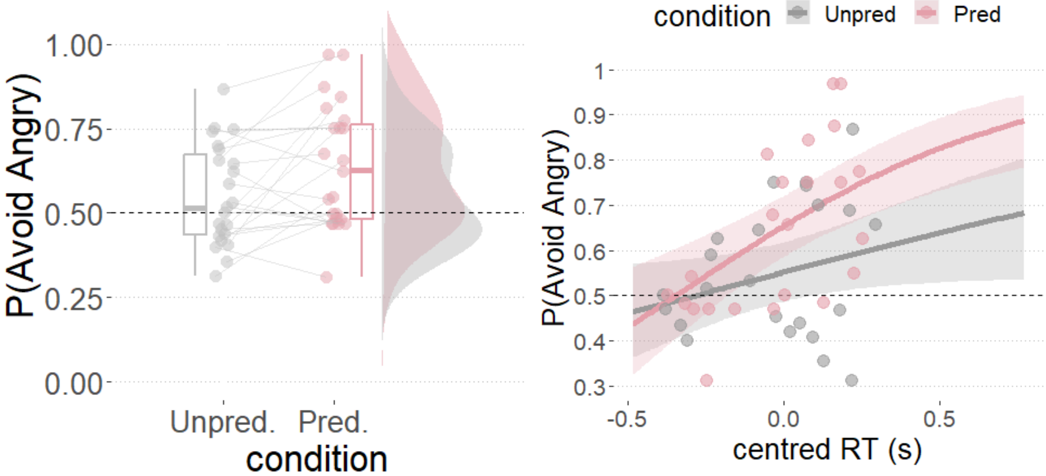

Figure A1*: A) Proportion of avoidance choices across conditions. B) Predicted probability of avoidance choices as a function of RTs. Points: observed avoidance rates over median RTs for each participant. Solid lines: group level predicted values. Shaded areas: 95% credible intervals. Dashed line: chance level (50%).*

B. Inclusion Criteria

Participants had to:

1. Be aged between 18 and 35 years old.
2. Be able to speak French fluently.
3. Not have major vision or hearing problems.
4. Not wear glasses or wear contact lenses on the day of the study.
5. Not be pregnant.
6. Not be claustrophobic.
7. Not suffer from postural instability or vestibular disorders.
8. Not suffer from motion sickness.
9. Not suffer from oculomotor disorders (strabismus, convergence disorder, anisophoria, anisometropia (asymmetrical refraction disorder), relative amblyopia (difference in visual perception between the two eyes), uncorrected ametropia, ocular pathologies or anomalies.
10. Not suffer from migraines.
11. Not have any affective disorders (depression, anxiety disorders, etc.) or behavioral disorders (anorexia, bulimia, addiction, etc.).
12. Have no medical history of neurological or psychiatric problems (epilepsy, depression, schizophrenia, apraxia, head trauma, etc.).
13. Not take any neurological, cardiac, or psychiatric medication.
14. Be willing to not consume alcohol the night before the study.
15. Be willing to not take any narcotics in the week preceding the study (marijuana, cocaine, ecstasy, MDMA, ketamine, etc.).
16. Agree to their data being made available anonymously to the scientific community criteria.

Furthermore, participants were asked not to drink alcohol the night before the experiment and not to assume drugs the week before the experiment.

From Experiment 2, participants were also required not to suffer from any cardiovascular disorders and not to have smoked tobacco for at least two hours before the experiment.

C. Stimuli Creation

Selection and Morphing of Facial Emotions

20 actors (10 females) have been selected from the Radboud Face Database (Langner et al., 2010). For each, two expressions have been selected (neutral and angry). Thus, the selected faces consisted of 10 actors x 2 genders x 2 emotions (N_{total} = 40). Later, two actors have been discarded due to their bad quality in virtual reality, leaving a final amount of 36 faces.

In order to create 3D head models from the 2D face stimuli, we used the DECA model (Detailed Expression Capture and Animation; Feng et al., 2021). The model projects a 2D image to a 3D model surface by "UV mapping". Spurious features of the UV maps thus obtained have been removed using a custom black mask. The Meshlab software (Cignoni et al., 2008) has then been used to align the 3D reconstructed morphs for stimuli display. Vertical alignment was made by keeping noses in line across actors and expressions, while horizontal alignment was made by keeping the eyes in line. The aligned 3D morphs have then been exported to the Blender software (Blender Foundation), where an eye-bulge correction has been performed using the "mesh" and "UV Map" editing tools in case of spurious reconstructions.

3D face-body merge

Four 3D human body models (two males, two females) presenting grey scale colored clothes and simple hairstyles have been purchased from the RenderPeople website. The validated 3D stimuli previously described have been merged with these models in Blender using the "slicing", "object joining" and "vertex stitching" editing tools. We also used the "smoothing" tool to sculpt the uneven vertices and edges in order to frame a realistic body with the new face. A texture color correction has been conducted to ensure smooth transition between body and face colors. Then, we created meta-rigs, i.e. a simulation of bones and articulation for each body, in order to control the posture in the Virtual Environment. Finally, we added a basic walk animation for the individuals entering and exiting the elevators. Such idle movements, consisting in small body oscillations, provide a more naturalistic view of the individuals (Treal et al., 2021).

D. Bayesian hierarchical GLM model fitting and selection

Model fitting

Bayesian Hierarchical Logistic Generalised Linear Models (or Generalized Linear Mixed Effect Model or GLMM) have been fitted using the R package *brms* (Bürkner, 2017). The package performs parameter estimates using a Markov Chain Monte Carlo (MCMC) sampling. In all experiments and for each model three chains were run in parallel with 10000 iterations each, 2000 of which have been used as a warm-up period. In the most complex model we fitted single-trial avoidance responses as a function of the Condition (Unpredictable = 0 vs Predictable = 1), the RTs (grand average centered) and their interaction. Individual participants have been introduced as a random effect keeping a maximal random slope structure. We used the following priors to fit the model:

β_0	~	N(0, 2.5)
$\beta_{predictors}$	~	N(0, 2.5)
cor	~	lkj(2)
SD	~	exponential(2)

Where β_0 is the prior for the intercept, $\beta_{predictors}$ is the prior for the slope of predictors, *N* is the gaussian distribution, *cor* is the correlation parameter, *lkg* is the Lewandowski-Kurowicka-Joe distribution.

Experiment 1

Model selection

Stepwise model selection has been performed by running several models with an increasing number of predictors and comparing their WAIC indexes. For Experiment 1 the best fitting model has been found to be the one including the Condition and RTs effect as predictors, but not their interaction (*Table D1*).

Table D1: *Regression Model comparison for Experiment 1.*

model	WAIC	SE
Inter	4361.40	52.70
Inter + Cond	4300.02	53.95
Inter + RTs	4333.41	53.74
Inter + Cond + RTs	**4287.98**	**54.66**

model	WAIC	SE
Inter + Cond + RTs + Cond:RTs	4289.75	54.75

Note: Bold: Best fitting model (i.e. lowest WAIC). Inter: Intercept. Cond: Condition. RTs: Response Times. SE: Standard Error

Model diagnostic

Chain Convergence has been visually inspected and verified. All chains presented a 'hairy caterpillar' shape. In *Table D2* and *D3* we provide R-hat and Effective Sample Size (ESS) statistics showing that chains properly converged and mixed.

Table D2: *Experiment 1 model results and diagnostic for each parameter.*

	Estimate	Est.Error	95%CrI-low	95%CrI-Up	$p_{>0}$	$p_{<0}$	BF_{01}	BF_{10}	R-hat	ESS
Inter	0.714	0.098	0.525	0.912	1	0	0.000	5.021e+14	1	11990
Cond	0.427	0.114	0.207	0.655	1	0	0.037	27.167	1	17005
RTs	1.173	0.222	0.745	1.618	1	0	0.000	1.105e+15	1	16197

Note: Inter: Intercept. Cond: Condition. RTs: Response Times. Est.Error: Estimate Error. 95%CrI-low and -up: lower and upper boundary of the 95% Credible Interval. $p_{>0}$ and $p_{<0}$: proportion of posterior samples greater and lower than 0. BF_{01} and BF_{10}: Bayes Factor for the parameter to be equal or different from 0. Rhat: R-hat potential scale reduction factor. ESS: Effective Sample Size.

Table D3*: Experiment 1 model random effect table*

	Estimate	Est.Error	95%CrI-low	95%CrI-up	R-hat	ESS
sd(Inter)	.493	.179	.065	.800	1.003	1048
sd(CondUnpred)	.324	.192	.016	.692	1.003	1184
sd(CondPred)	.505	.192	.069	.849	1.003	909
sd(RTs)	.649	.367	.034	1.363	1.001	3337
cor(Inter,CondUnpred)	-.054	.370	-.690	.675	1.000	3889
cor(Inter,CondPred)	.060	.313	-.500	.699	1.000	3294
cor(CondUnpred,CondPred)	.059	.372	-.673	.689	1.003	1302
cor(Inter,RTs)	.100	.333	-.561	.714	1.000	8774
cor(CondUnpred,RTs)	.079	.352	-.624	.708	1.000	7875
cor(CondPred,RTs)	-.014	.325	-.629	.636	1.001	10508

Note: sd: Standard Deviation. cor: correlations. Inter: Intercept. CondUnpred and -Pred : Unpredictable and Predictable Condition. Est.Error: Estimate Error. 95%CrI-low and -up: lower and upper boundary of the 95% Credible Interval. Rhat: R-hat potential scale reduction factor. ESS: Effective Sample Size.

Experiment 2

Model selection

For Experiment 2 the best fitting model was the one including the Condition and RTs effect as predictors, as well as their interaction (*Table D4*).

Table D4: *Regression Model comparison for Experiment 2.*

model	WAIC	SE
Inter	2428.00	26.72
Inter + Cond	2393.97	28.62
Inter + RTs	2413.46	27.64
Inter + Cond + RTs	2385.47	28.99
Inter + Cond + RTs + Cond:RTs	**2380.13**	**29.38**

Note: Inter: Intercept. Cond: Condition. RTs: Response Times. Bold: Best fitting model (i.e. lowest WAIC). SE: Standard Error

Model diagnostic

Chain Convergence has been visually inspected and verified. All chains presented a 'hairy caterpillar' shape. In *Tables D5* and *D6* we provide R-hat and Effective Sample Size (ESS) statistics showing that chains properly converged and mixed.

Table D5: *Experiment 2 model results and diagnostic for each parameter.*

	Estimate	Est.Error	95%CrI-low	95%CrI-up	$p_{>0}$	$p_{<0}$	BF_{01}	BF_{10}	R-hat	ESS
Inter	.427	.121	.192	.670	1	0	.068	14.610	1	17386
Cond	.285	.137	.012	.553	.979	.021	2.203	.454	1	19527
RTs	.473	.312	-.132	1.091	.937	.063	2.473	.404	1	22328
Cond:RTs	1.470	.392	.681	2.219	1	0	.010	102.610	1	26827

Note: Inter: Intercept. Cond: Condition. RTs: Response Times. Est.Error: Estimate Error. 95%CrI-low and -up: lower and upper boundary of the 95% Credible Interval. $p_{>0}$ and $p_{<0}$: Proportion of posterior samples greater and lower than 0. BF_{01} and BF_{10}: Bayes Factor for the parameter to be equal or different from 0. Rhat: R-hat potential scale reduction factor. ESS: Effective Sample Size.

Table D6: *Experiment 2 model random effect table*

	Estimate	Est.Error	95%CrI-low	95%CrI-up	R-hat	ESS
sd(Inter)	.235	.137	.014	.519	1.001	5849
sd(CondUnpred)	.412	.160	.068	.719	1.001	4637

	Estimate	Est.Error	95%CrI-low	95%CrI-up	R-hat	ESS
sd(CondPred)	.199	.141	.007	.519	1	6340
sd(RTs)	.683	.330	.057	1.335	1.002	7733
sd(CondPred:RTs)	.378	.319	.011	1.166	1	13868
cor(Inter,CondUnpred)	.048	.331	-.572	.677	1	8546
cor(Inter,CondPred)	-.068	.355	-.704	.629	1	22224
cor(CondUnpred,CondPred)	.032	.340	-.639	.654	1	19417
cor(Inter,RTs)	.235	.333	-.484	.785	1	9339
cor(CondUnpred,RTs)	.293	.312	-.399	.799	1	13902
cor(CondPred,RTs)	.047	.345	-.621	.687	1	16222
cor(Inter,CondPred:RTs)	.036	.349	-.637	.679	1	34607
cor(CondUnpred,CondPred:RTs)	-.070	.349	-.700	.616	1	31041
cor(CondUnpred,CondPred:RTs)	.070	.355	-.624	.701	1	25253
cor(RTs,CondPred:RTs)	-.058	.352	-.693	-.638	1	27832

Note: sd: Standard Deviation. cor: Correlation. Inter: Intercept. Cond: Condition. RTs: Response Times. Est.Error: Estimate Error. 95%CrI-low and -up: lower and upper boundary of the 95% Credible Interval. Rhat: R-hat potential scale reduction factor. ESS: Effective Sample Size.

Experiment 3

Model Selection

For Experiment 3 the best fitting model was the one including the Condition and RTs effect as predictors, as well as their interaction (*Table D7*).

Table D7: *Regression Model comparison for Experiment 3.*

model	WAIC	SE
Inter	4929.67	34.17
Inter + Cond	4894.62	35.90
Inter + RTs	4911.34	35.28
Inter + Cond + RTs	4878.56	36.70
Inter + Cond + RTs + Cond:RTs	**4871.64**	**37.02**

Note: Inter: Intercept. Cond: Condition. RTs: Response Times. Bold: Best fitting model (i.e. lowest WAIC). SE: Standard Error

Model diagnostic

Chain Convergence has been visually inspected and verified. All chains presented a 'hairy caterpillar' shape. In *Table D8* and *D9* we provide R-hat and Effective Sample Size (ESS) statistics showing that chains properly converged and mixed.

Table D8: *Experiment 3 model results and diagnostic for each parameter.*

	Estimate	Est.Error	95%CrI-low	95%CrI-Up	$p_{>0}$	$p_{<0}$	BF_{01}	BF_{10}	R-hat	ESS
Inter	.306	.080	.150	.465	1	0	.073	13.712	1	16607
Cond	.187	.091	.009	.369	.980	.020	3.423	.292	1	19213
RTs	.716	.245	.236	1.197	.998	.002	.151	6.636	1	18952
Cond:RTs	1.120	.330	.463	1.772	.999	.001	.039	25.851	1	21753

Note: *Inter: Intercept. Cond: Condition. RTs: Response Times. Est.Error: Estimate Error. 95%CrI-low and -up: lower and upper boundary of the 95% Credible Interval. $p_{>0}$ and $p_{<0}$: proportion of posterior samples greater and lower than 0. BF_{01} and BF_{10}: Bayes Factor for the parameter to be equal or different from 0. Rhat: R-hat potential scale reduction factor. ESS: Effective Sample Size.*

Table D9: *Experiment 3 model random effect table*

	Estimate	Est.Error	95%CrI-low	95%CrI-up	R-hat	ESS
sd(Inter)	.327	.122	.042	.541	1	2355
sd(CondUnpred)	.375	.145	.019	.553	1.001	2150
sd(CondPred)	.237	.141	.012	.517	1	2711
sd(RTs)	.683	.300	.080	1.259	1.001	6335
sd(CondPred:RTs)	.941	.482	.058	1.861	1.002	3392
cor(Inter,CondUnpred)	.010	.329	-.588	.650	1.001	9748
cor(Inter,CondPred)	-.039	.338	-.643	.634	1	10096
cor(CondUnpred,CondPred)	.058	.344	-.629	.670	1.001	5585
cor(Inter,RTs)	.439	.297	-.286	.868	1	4038
cor(CondUnpred,RTs)	.244	.331	-.460	.797	1	7401
cor(CondPred,RTs)	.103	.339	-.577	.710	1	6812
cor(inter,CondPred:RTs)	-.052	.320	-.645	.586	1	8128
cor(CondUnpred,CondPred:RTs)	-.236	.321	-.770	.466	1.001	8169
cor(CondPred,CondPred:RTs)	.191	.332	-.510	.753	1	8654
cor(RTs,CondPred:RTs)	-.157	.322	-.701	.525	1	11114

Note: Est.Error: Estimate Error. 95%CrI-low and -up: lower and upper boundary of the 95% Credible Interval. R-hat: R-hat potential scale reduction factor. ESS: Effective Sample Size.

Experiment 3 – Effect of Heart Rate and Subjective Value on Avoidance Decisions

Chain Convergence has been visually inspected and verified. All chains presented a ‘hairy caterpillar’ shape. In *Table D10* we provide R-hat and Effective Sample Size (ESS) statistics showing that chains properly converged and mixed.

Table D10: *Regression table for the logistic model predicting avoidance by Condition, RTs, HR and SV*

	Est.	Est.Err.	95%CrI-low	95%CrI-Up	$p_{>0}$	$p_{<0}$	BF_{01}	BF_{10}	Rhat	ESS
Intercept	.183	.115	-0.043	0.408	.945	.055	6.269	0.160	1	13685
Cond	.322	.150	0.023	0.617	.982	.018	1.871	0.534	1	13306
HR	-.032	.106	-0.239	0.174	.380	.620	22.603	0.044	1	13106
RTs	.298	.104	0.095	0.501	.998	.002	0.454	2.200	1	15587
SV	.229	.240	-0.236	0.706	.829	.171	6.743	0.148	1	14802
Cond:HR	.250	.146	-0.037	0.536	.957	.043	3.888	0.257	1	13505
Cond:RTs	.189	.150	-0.101	0.487	.897	.103	7.793	0.128	1	15914
HR:RTs	-.105	.110	-0.322	0.109	.167	.833	14.226	0.070	1	12732
Cond:SV	-.131	.337	-0.786	0.541	.347	.653	6.732	0.149	1	13446
HR:SV	.171	.251	-0.319	0.664	.754	.246	8.168	0.122	1	12593
RTs:SV	-.291	.236	-0.758	0.165	.108	.892	5.027	0.199	1	13961
Cond:HR:RTs	.177	.156	-0.130	0.485	.873	.127	8.429	0.119	1	13496
Cond:HR:SV	-.946	.354	-1.644	-0.256	.004	.996	0.214	4.681	1	12906
Cond:RTs:SV	.467	.341	-0.203	1.130	.914	.086	3.060	0.327	1	15214
HR:RTs:SV	.515	.268	-0.006	1.051	.973	.027	1.484	0.674	1	12449
Cond:HR:RTs:SV	-.931	.380	-1.685	-0.191	.007	.993	0.309	3.236	1	12963

Note: Cond: Condition. HR: heart rate. RTs: response times. SV: subjective value. Est.: Estimate. Est.Error: Estimate Error. 95%CrI-low and -up: lower and upper boundary of the 95% Credible Interval. $p_{>0}$ and $p_{<0}$: Proportion of posterior samples greater and lower than 0. BF_{01} and BF_{10}: Bayes Factor for the parameter to be equal or different from 0. Rhat: R-hat potential scale reduction factor. ESS: Effective Sample Size.

Experiment 2 and 3 by class

Chain Convergence has been visually inspected and verified. All chains presented a 'hairy caterpillar' shape. In *Table D11* we provide R-hat and Effective Sample Size (ESS) statistics showing that chains properly converged and mixed.

Table D11: *Regression tables for logistic model including classes*

	Est.	Est.Err.	95%CrI-low	95%CrI-Up	$p_{>0}$	$p_{<0}$	BF_{01}	BF_{10}	Rhat	ESS
Inter	0.883	0.084	0.718	1.048	1.000	0.000	0.000	4.695965e+15	1	11960
Cond	0.102	0.123	-0.136	0.346	0.797	0.203	5.933	1.690000e-01	1	12154
RTs	1.618	0.250	1.125	2.105	1.000	0.000	0.000	9.043625e+17	1	14776
Class	-0.820	0.099	-1.015	-0.625	0.000	1.000	0.000	6.745862e+38	1	11761
Cond:RTs	0.503	0.364	-0.214	1.216	0.917	0.083	1.089	9.180000e-01	1	14720
Cond:Class	0.220	0.146	-0.069	0.505	0.934	0.066	2.173	4.600000e-01	1	11761
RTs:Class	-1.570	0.289	-2.135	-1.008	0.000	1.000	0.000	5.179647e+17	1	14924
Cond:RTs: Class	1.165	0.425	0.338	1.995	0.996	0.004	0.053	1.885800e+01	1	14764

Note: Inter: Intercept. Cond: Condition. RTs: Response Times. Est.: Estimate. Est.Err.: Estimate Error. 95%CrI-low and -up: lower and upper boundary of the 95% Credible Interval. $p_{>0}$ and $p_{<0}$: Proportion of posterior samples greater and lower than 0. BF_{01} and BF_{10}: Bayes Factor for the parameter to be equal or different from 0. Rhat: R-hat potential scale reduction factor. ESS: Effective Sample Size.

As a robustness check, we fitted separate Bayesian hierarchical logistic models for each class. We found that individuals in Class A exhibited a random avoidance rate in the Unpredictable condition (Intercept: 95%CrI=[-.042, .163]; BF_{01}=9.977), but a credibly higher rate in the Predictable condition (Condition effect: β=.326; 95%CrI=[.164, .491]; $p_{<0}$=.0; BF_{10}=60.712; *Figure 4C₂*). Furthermore, while we did not find a credible main effect of RTs in Class A (95%CrI=[-.301, .341]; BF_{01}=6.134), we did find a credible Condition by RTs interaction (β=1.629; 95%CrI=[1.097, 2.146]; $p_{<0}$=.0; BF_{10}=1.219e+16), showing that in Class A, slower RTs were associated with a higher probability of avoiding the angry avatar in the Predictable condition, but not in the Unpredictable one (*Figure 4D₂*). Conversely, Class B avoided the angry avatars more than chance throughout the task (Intercept: β=.900; 95%CrI=[.713, 1.099]; $p_{<0}$=.0; BF_{10}=2.421e+15) with no difference between conditions (Condition effect: 95%CrI=[-.158, .316], BF_{01}=6.547; *Figure 4C₁*). Moreover, we found a main effect of RTs on avoidance rate for this class (RTs: β=1.753; 95%CrI=[1.179, 2.318]; $p_{<0}$=.0; BF_{10}=1.960e+16), with slower RTs associated with a higher probability of avoidance regardless of the condition (Condition x RTs: 95%CrI=[-.497, 1.005], BF_{01}=2.121; *Figure 4D₁*).

Table D12: *Regression tables for separate logistic models in Class A and B*

	Est.	Est.Err.	95%CrI-low	95%CrI-Up	$p_{>0}$	$p_{<0}$	BF_{01}	BF_{10}	Rhat	ESS
					Class A					
Inter	0.059	0.052	-0.042	0.163	0.874	0.126	9.977	1.000000e-01	1	17808
Cond	0.326	0.083	0.164	0.491	1.000	0.000	0.016	6.071200e+01	1	21925
RTs	0.020	0.164	-0.301	0.341	0.548	0.452	6.134	1.630000e-01	1	22049
Cond:RTs	1.629	0.267	1.097	2.146	1.000	0.000	0.000	1.218985e+16	1	27715
					Class B					
Inter	0.900	0.098	0.713	1.099	1.000	0.000	0.000	2.420635e+15	1	17502
Cond	0.083	0.121	-0.158	0.316	0.757	0.243	6.547	1.530000e-01	1	14901
RTs	1.753	0.291	1.179	2.318	1.000	0.000	0.000	1.959281e+16	1	16838
Cond:RTs	0.249	0.385	-0.497	1.005	0.739	0.261	2.121	4.710000e-01	1	13549

Note: Inter: Intercept. Cond: Condition. RTs: Response Times. Est.: Estimate. Est.Err.: Estimate Error. 95%CrI-low and -up: lower and upper boundary of the 95% Credible Interval. $p_{>0}$ and $p_{<0}$: Proportion of posterior samples greater and lower than 0. BF_{01} and BF_{10}: Bayes Factor for the parameter to be equal or different from 0. Rhat: R-hat potential scale reduction factor. ESS: Effective Sample Size.

E. DDM model fitting and selection

Model fitting

In all experiments Bayesian Hierarchical Drift Diffusion Models have been fitted using the Python package *HDDM* (Wiecki et al., 2013). The package performs parameter estimates using Markov Chain Monte Carlo (MCMC) sampling. For each model three chains were run in parallel with 20000 iterations each (5000 of which as burn-in period) and a thinning factor of 3. We used the population-informed priors by Matzke and Wagenmakers (2009) to fit the models. In all models the starting point bias has been kept fixed and equidistant between boundaries.

Experiment 1

Model Selection

We compared several models by sequentially allowing the drift rate, the boundary separation, the non-decision time, combinations of these or none of them to vary across conditions. In this experiment all criteria favored as best fitting the model allowing the drift rate and the non-decision time to vary across conditions (*Table E1*).

Table E1: *Experiment 1 DDM model selection.*

model	DIC	BPIC
Cond(a,t0,v)	**2896.115**	**3055.893**
Cond(a,t0)+v	2925.287	3081.306
Cond(a,v)+t0	2901.317	3066.291
Cond(v+t0)+a	2899.123	3056.410
Cond(a)+t0+v	2931.862	3095.850
Cond(t0)+a+v	2924.362	3077.996
Cond(v)+a+t0	2912.539	3076.557
a+t0+v	2938.374	3101.315

Note: Cond: Condition, between parentheses the parameter(s) allowed to vary. a: Boundary separation. v: Drift rate. t0: Non-decision time. DIC: Deviance Information Criterion. BPIC: Bayesian Predictive Information Criterion. Bold: best fitting model.

Model Diagnostic

Chain Convergence has been visually inspected and verified. All chains presented a ‘hairy caterpillar’ shape. In *Table E2* and *E3* we provide R-hat and Effective Sample Size (ESS) statistics showing that chains properly converged and mixed.

Table E2: *Experiment 1 best fitting model group-level mu parameters.*

parameter	Estimate	95%CrI-low	95%CrI-up	$p_{>0}$	$p_{<0}$	R-hat	ESS
a (Inter)	1.219	1.145	1.298	1	0	1	11142
a (Cond)	.048	.009	.089	.990	.010	1	5022
v (Inter)	.575	.398	.753	1	0	1	12315
v (Cond)	.344	.226	.460	1	0	1	10385
t0 (Inter)	.380	.336	.436	1	0	1	12311
t0 (Cond)	.005	-.002	.012	.915	.085	1	3590

Note: Inter: Intercept. Cond: Condition. a: Boundary Separation. t0: Non-decision time. v: Drift rate. Estimate: median posterior estimate. 95%CrI-low and -up: Lower and upper boundaries of the 95% Credible Interval. $p_{>0}$ and $p_{<0}$: Proportion of posterior samples greater and lower than zero. R-hat: R-hat potential scale reduction factor. ESS: Effective Sample Size.

Table E3: *Experiment 1 best fitting model group-level sd parameters*

parameter	Estimate	95%CrI-low	95%CrI-Up	$p_{>0}$	$p_{<0}$	R-hat	ESS
sd (a Inter)	.274	.225	.342	1	0	1	12121
sd(v Inter)	.615	.495	.773	1	0	1	14181
sd(t0 Inter)	.193	.157	.245	1	0	1	13065

Note: Inter: Intercept. a: Boundary separation. t0: Non-decision time. v: Drift rate. sd: Standard deviation. Estimate: median posterior estimate. 95%CrI-low and -up: Lower and upper boundaries of the 95% Credible Interval. $p_{>0}$ and $p_{<0}$: Proportion of posterior samples greater and lower than zero. R-hat: R-hat potential scale reduction factor. ESS: Effective Sample Size.

Posterior Predictive Check

In order to assess the reliability of our model we simulated data from the best fitting parameters and we assessed their resemblance to the observed data. Simulated and Observed data present high correlations for avoidance response proportion and median RTs for avoidance and approach responses (*Figure E1A-C*).

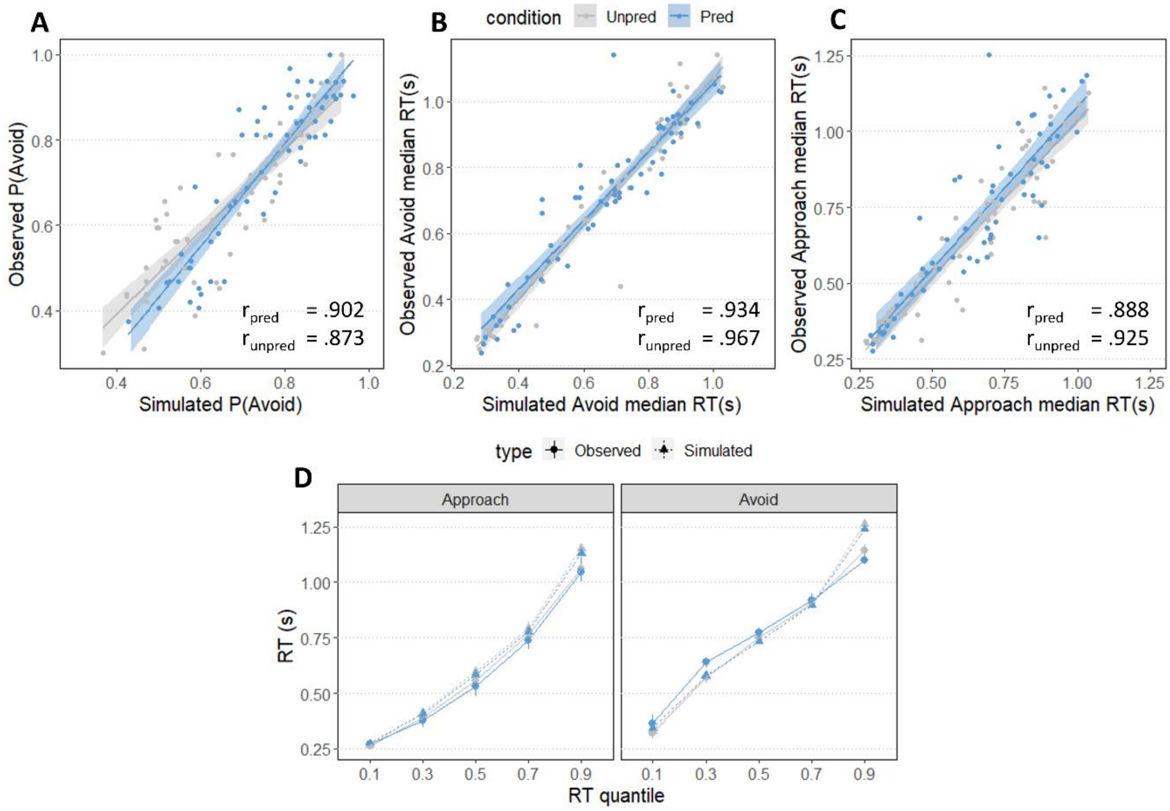

Figure E1: *Posterior predictive check for the best fitting DDM in Experiment 1. Top row: correlation between observed and simulated proportion of avoidance responses (A), median RTs for avoidance responses (B) and median RTs for approach responses (C). D) Observed and Simulated RTs by quantile of the RTs distribution parameters.*

Experiment 2

Model Selection

In Experiment 2 all criteria favored as best fitting the model allowing the drift rate and the non-decision time to vary by condition (*Table E4*).

Table E4: *Experiment 2 DDM model selection.*

model	DIC	BPIC
Cond(a,t0,v)	1390.345	1466.925
Cond(a,t0)+v	1401.504	1477.468
Cond(a,v)+t0	1394.436	1473.379
Cond(v+t0)+a	**1387.701**	**1462.210**
Cond(a)+t0+v	1405.457	1483.619
Cond(t0)+a+v	1400.267	1476.016
Cond(v)+a+t0	1393.391	1471.016

model	DIC	BPIC
a+t0+v	1398.445	1470.497

Note: Cond: Condition, between parentheses the parameter(s) allowed to vary. a: Boundary separation. v: Drift rate. t0: Non-decision time. DIC: Deviance Information Criterion. BPIC: Bayesian Predictive Information Criterion. Bold: best fitting model.

Model Diagnostic

Chain Convergence has been visually inspected and verified. All chains presented a 'hairy caterpillar' shape. In *Table E5* and *E6* we provide R-hat and Effective Sample Size (ESS) statistics showing that chains properly converged and mixed.

Table E5: *Experiment 2 best fitting model group-level mu parameters.*

parameter	Estimate	95%CrI-low	95%CrI-up	$p_{>0}$	$p_{<0}$	R-hat	ESS
a	1.150	1.042	1.274	1	0	1	12102
v (Inter)	.384	.161	.605	1	0	1	11339
v (Cond)	.303	.134	.472	1	0	1	10370
t0 (Inter)	.370	.304	.461	1	0	1	10790
t0 (Cond)	.007	0	.015	.967	.033	1	2533

Note: Inter: Intercept. Cond: Condition. a: Boundary Separation. t0: Non-decision time. v: Drift rate. Estimate: median posterior estimate. 95%CrI-low and -up: Lower and upper boundary of the 95% Credible Interval. $p_{>0}$ and $p_{<0}$: Proportion of posterior samples greater and lower than zero. R-hat: R-hat potential scale reduction factor. ESS: Effective Sample Size.

Table E6: *Experiment 2 best fitting model group-level sd parameters.*

parameter	Estimate	95%CrI-low	95%CrI-Up	$p_{>0}$	$p_{<0}$	R-hat	ESS
sd(a Inter)	.299	.228	.415	1	0	1	12359
sd(t0 Inter)	.490	.344	.694	1	0	1	11031
sd(v Inter)	.213	.159	.309	1	0	1	10451

Note: Inter: Intercept. a: Boundary separation. t0: Non-decision time. v: Drift rate. Pred and Unpred: Predictable and Unpredictable conditions. sd: Standard deviation. cor: correlation. Estimate: median posterior estimate. 95%CrI-low and -up: Lower and upper boundary of the 95% Credible Interval. $p_{>0}$ and $p_{<0}$: Proportion of posterior samples greater and lower than zero. R-hat: R-hat potential scale reduction factor. IAT: Integrated Autocorrelation Time. ESS: Effective Sample Size.

Posterior predictive check

In order to assess the reliability of our model we simulated data from the best fitting parameters and we assessed their resemblance to the observed data. While the observed and simulated RTs show a high correlation for both avoidance and approach choices and for both the Predictable and Unpredictable conditions (*Figure E2-B and C*), the correlation of the proportion of avoidance responses between is less good (*Figure E2-A*). This is likely due to the facts that: (a) the task contained only a limited number of trials (issue that in Experiment 1 and 3 was balanced by a double number of participants); (b) the sample in Experiment 2 presented latent classes presenting different behavioral patterns, as illustrated by finite mixture modeling.

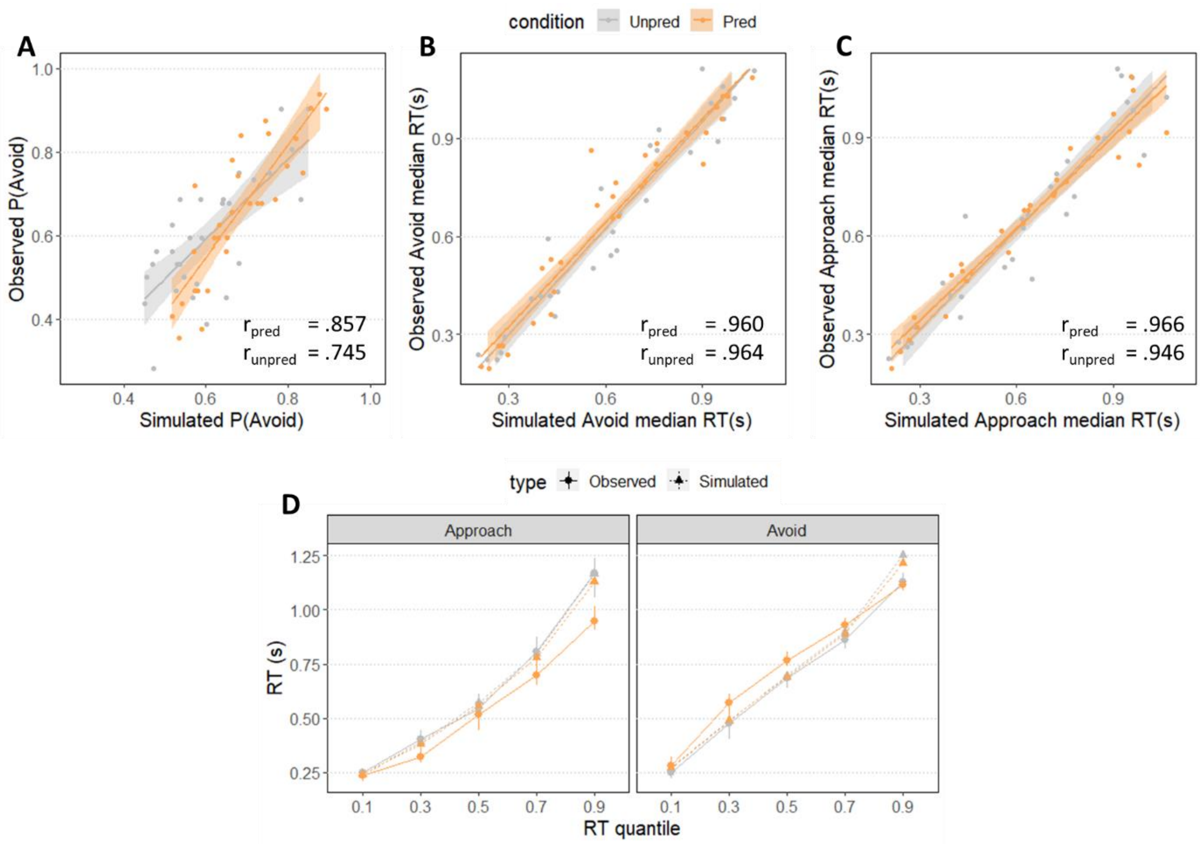

Figure E2: *Posterior predictive check for the best fitting DDM in Experiment 2. Top row: correlation between observed and simulated proportion of avoidance responses (A), median RTs for avoidance responses (B) and median RTs for approach responses (C). D) Observed and Simulated RTs by quantile of the RTs distribution.*

Experiment 3

Model Selection

In Experiment 2 all criteria favored as best fitting the model allowing the drift rate and the non-decision time to vary across conditions (*Table E7*).

Table E7: *Experiment 3 DDM model selection.*

model	DIC	BPIC
Cond(a,t0,v)	**1744.723**	**1890.045**
Cond(a,t0)+v	1750.789	1895.402
Cond(a,v)+t0	1749.392	1882.588
Cond(v+t0)+a	1745.979	1888.620
Cond(a)+t0+v	1759.447	1896.433
Cond(t0)+a+v	1753.228	1896.395
Cond(v)+a+t0	1755.303	1895.344
a+t0+v	1758.637	1894.979

Note: Cond: Condition, between parentheses the parameter(s) allowed to vary. a: Boundary separation. v: Drift rate. t0: Non-decision time. DIC: Deviance Information Criterion. BPIC: Bayesian Predictive Information Criterion. Bold: best fitting model.

Model Diagnostic

Chain Convergence has been visually inspected and verified. All chains presented a 'hairy caterpillar' shape. In *Table E8* and *E9* we provide R-hat and Effective Sample Size (ESS) statistics showing that chains properly converged and mixed.

Table E8: *Experiment 3 best fitting model group-level mu parameters.*

parameter	Estimate	95%CrI-low	95%CrI-up	$p_{>0}$	$p_{<0}$	R-hat	ESS
a (Inter)	1.049	.986	1.117	1	0	1	12376
a (Cond)	.035	.002	.068	.981	.019	1	4936
t0 (Inter)	.340	.297	.393	1	0	1	11841
t0 (Cond)	-.011	-.017	-.006	0	1	1	3349
v (Inter)	.298	.131	.468	.999	.001	1	9067
v (Cond)	.185	.056	.312	.997	.003	1	9946

Note: Inter: Intercept. Cond: Condition. a: Boundary Separation. t0: Non-decision time. v: Drift rate. Estimate: median posterior estimate. 95%CrI-low and -up: Lower and upper boundaries of the 95% Credible Interval. $p_{>0}$ and $p_{<0}$: Proportion of posterior samples greater or lower than zero. R-hat: R-hat potential scale reduction factor. ESS: Effective Sample Size.

Table E9: *Experiment 2 best fitting model group-level sd parameters*

parameter	Estimate	95%CrI-low	95%CrI-Up	$p_{>0}$	$p_{<0}$	R-hat	ESS
sd (a Inter)	.240	.196	.299	1	0	1	11313

parameter	Estimate	95%CrI-low	95%CrI-Up	$p_{>0}$	$p_{<0}$	R-hat	ESS
sd(t0 Inter)	.182	.147	.234	1	0	1	12479
sd(v Inter)	.543	.432	.687	1	0	1	11297

Note: a: Boundary separation. t0: Non-decision time. v: Drift rate. sd: Standard deviation. Estimate: median posterior estimate. 95%CrI-low and -up: Lower and upper boundary of the 95% Credible Interval. $p_{>0}$ and $p_{<0}$: Proportion of posterior samples greater or lower than zero. R-hat: R-hat potential scale reduction factor. ESS: Effective Sample Size.

Posterior predictive check

In order to assess the reliability of our model we simulated data from the best fitting parameters and we assessed their resemblance to the observed data. Simulated and Observed data present high correlations for avoidance response proportion and median RTs for avoidance and approach responses (*Figure E3A-C*).

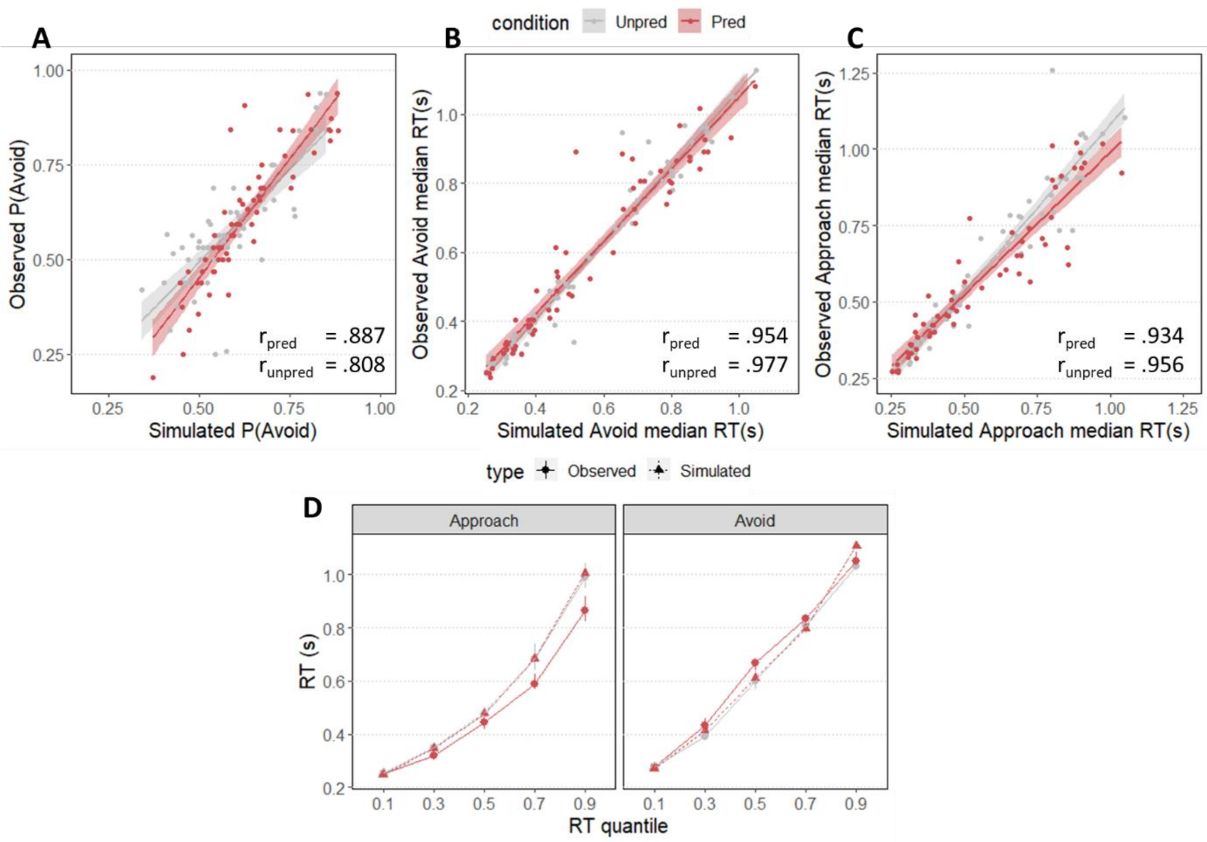

Figure E3: *Posterior predictive check for the best fitting DDM in Experiment 3. Top row: correlation between observed and simulated proportion of avoidance responses (A), median RTs for avoidance responses (B) and median RTs for approach responses (C). D) Observed and Simulated RTs by quantile of the RTs distribution.*

Experiment 2 and 3 by class

Model Selection

In Experiment 2 all criteria favored as best fitting the model allowing the drift rate and the non-decision time to vary across conditions (*Table E10*).

Table E10: *DDM model selection including latent classes*

model	DIC	BPIC
Class x Cond(a,t0,v)	2831.372	3211.618
Class x Cond(a,t0)+v	2934.104	3266.796
Class x Cond(a,v)+t0	3075.491	3387.706
Class x Cond(v+t0)+a	2781.559	3082.148
Class x Cond(a)+t0+v	3178.824	3445.030
Class x Cond(t0)+a+v	3150.556	3370.532
Class x Cond(v)+a+t0	3058.893	3296.496
Class(a+t0+v)	3149.917	3351.464

Note: Cond: Condition, between parentheses the parameter(s) allowed to vary. a: Boundary separation. v: Drift rate. t0: Non-decision time. DIC: Deviance Information Criterion. BPIC: Bayesian Predictive Information Criterion. Bold: best fitting model.

Model Diagnostic

Chain Convergence has been visually inspected and verified. All chains presented a ‘hairy caterpillar’ shape. In *Table E11* and *E12* we provide R-hat and Effective Sample Size (ESS) statistics showing that chains properly converged and mixed.

Table E11: *Best fitting model group-level mu parameters.*

parameter	Estimate	95%CrI-low	95%CrI-up	$p_{>0}$	$p_{<0}$	R-hat	ESS
class A: a	1.163	1.067	1.261	1	0	1	12913
classB: a	1.053	.988	1.121	1	0	1	13626
class A: v Unpred	.898	.735	1.068	1	0	1	8196
classB: v Unpred	.061	-.059	.178	.839	.161	1	9234
class A: v Pred	1.109	.850	1.370	1	0	1	13902
class B: v Pred	.303	.123	.476	1	0	1	13924
class A: t0 Unpred	.380	.323	.440	1	0	1	13960
classB: t0 Unpred	.325	.286	.371	1	0	1	13293
class A: t0 Pred	.410	.348	.475	1	0	1	13887
class B: t0 Pred	.330	.288	.375	1	0	1	13369

Note: a: Boundary Separation. t0: Non-decision time. v: Drift rate. Estimate: median posterior estimate. 95%CrI-low and -up: Lower and upper boundaries of the 95% Credible Interval. $p_{>0}$ and $p_{<0}$: Proportion of posterior samples greater or lower than zero. R-hat: R-hat potential scale reduction factor. ESS: Effective Sample Size.

Table E12: *Best fitting model group-level sd parameters*

parameter	Estimate	95%CrI-low	95%CrI-Up	$p_{>0}$	$p_{<0}$	R-hat	ESS
sd (a)	.255	.217	.304	1	0	1	13007
sd (v Unpred)	.292	.172	.414	1	0	1.001	2027
sd(v Pred)	.589	.473	.725	1	0	1	9579
sd (t0 Unpred)	.181	.152	.219	1	0	1	12766
sd(t0 Pred)	.188	.159	.228	1	0	1	12257

Note: a: Boundary separation. t0: Non-decision time. v: Drift rate. sd: Standard deviation. Estimate: median posterior estimate. 95%CrI-low and -up: Lower and upper boundary of the 95% Credible Interval. $p_{>0}$ and $p_{<0}$: Proportion of posterior samples greater or lower than zero. R-hat: R-hat potential scale reduction factor. ESS: Effective Sample Size. Unpred and Pred: Unpredictable and Predictable condition.

Posterior predictive check

In order to assess the reliability of our model we simulated data from the best fitting parameters and we assessed their resemblance to the observed data. Simulated and Observed data present high correlations for avoidance response proportion and median RTs for avoidance and approach responses (*Figure E4A-C*).

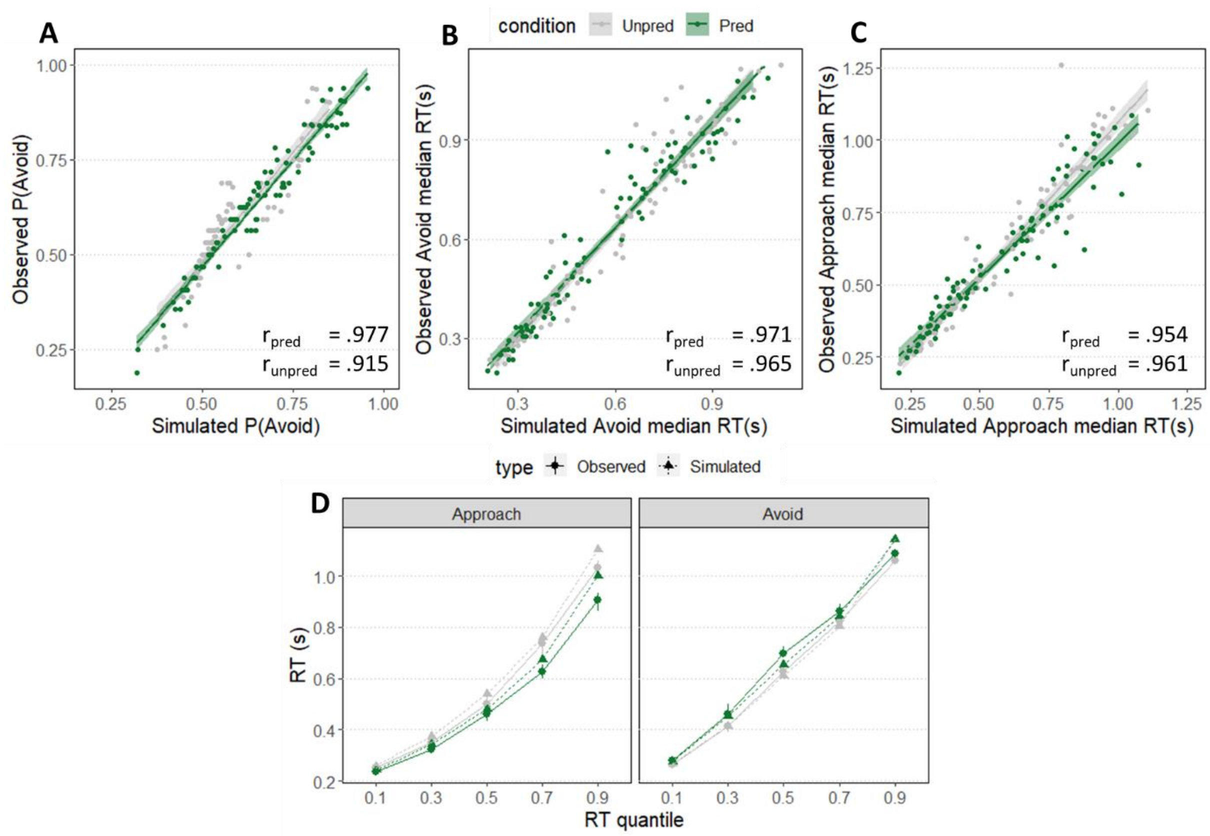

Figure E4: *Posterior predictive check for the best fitting DDM including latent classes. Top row: correlation between observed and simulated proportion of avoidance responses (A), median RTs for avoidance responses (B) and median RTs for approach responses (C). D) Observed and Simulated RTs by quantile of the RTs distribution.*

F. Head-Tracking

The Unity-UXF interface (Brookes et al., 2020) provides a continuous measure of head rotation for each trial on the x, y and z axis (i.e., pitch, roll and yaw rotation values) sampled at 60 frames per second. The rotation measurement represents participants' fixation (i.e., values of x and z are 0 if participants orient their head in front of them with a 90° angle from their torso) regardless of the position of the body in the virtual environment. In order to explore participants' visual exploration in relation to avatars' faces during the trial, we computed a continuous measure of rotation toward or away the face of the avatar in the chosen elevator in a given trial.

Data preprocessing and analysis

Initially, we define a 'fixation point' within the 3D environment by the x and z coordinates of the camera (i.e. the moving participants' position over time) and the y coordinate of the avatar's face in the chosen elevator. The 'distance of fixation' was then defined as the horizontal distance between the camera and this fixation point. Subsequently, we employed the tangent function to calculate the frame-by-frame displacement of the fixation point from its original position depending on participants' head orientation on the horizontal and vertical axes. This computation allowed for a continuous update of the fixation point. Next, we computed three sides of a two-dimensional triangle within a three-dimensional space. These sides included the Euclidean distances between the avatar's face in the chosen elevator, the fixation point, and the camera. With these distances, we could then calculate frame-by-frame the angle between the participants' fixation and the avatar's face using Carnot's theorem. Finally, since head rotation from avatars varied also depending on the camera's automatic movement in space (i.e. when entering the elevator during the outcome phase), we adjusted these frame-by-frame measures by subtracting the hypothetical angle that participants would have presented assuming no movement. This adjustment yielded an actual measure of head movement in degrees, indicating whether participants moved their head away (positive values) or toward (negative values) the face in the chosen elevator.

The resulting head rotation time series were analyzed using the one-dimensional statistical parametric mapping approach used for physiological data in Experiment 3 and described in the main text of the manuscript (*Experiment 3 - Material and Method - Data Analysis - Physiological data*). Before fitting the GLM, data were z-scored within participants using mean and standard deviation of the concatenated trial data.

Results

In all experiment we found a significant negative cluster for the intercept at the end of the trial during the outcome phase (Intercept - Experiment 1: t_{mass}=-3131.9, p_{cor}= 0; Experiment 2: t_{exp2}=-1427.3; p_{exp2}=.032; Experiment 3: t_{mass}=-2134.2; p_{corr}=.001; *Figure F1*). A negative cluster indicated that participants reduced the angle of head orientation from the face in the chosen elevator, i.e., they looked at the avatar. This result shows that regardless of the Outcome of the choice, the Condition and the RTs, participants oriented their head (and likely their attention) toward the face of the avatar, thus receiving the intended feedback.

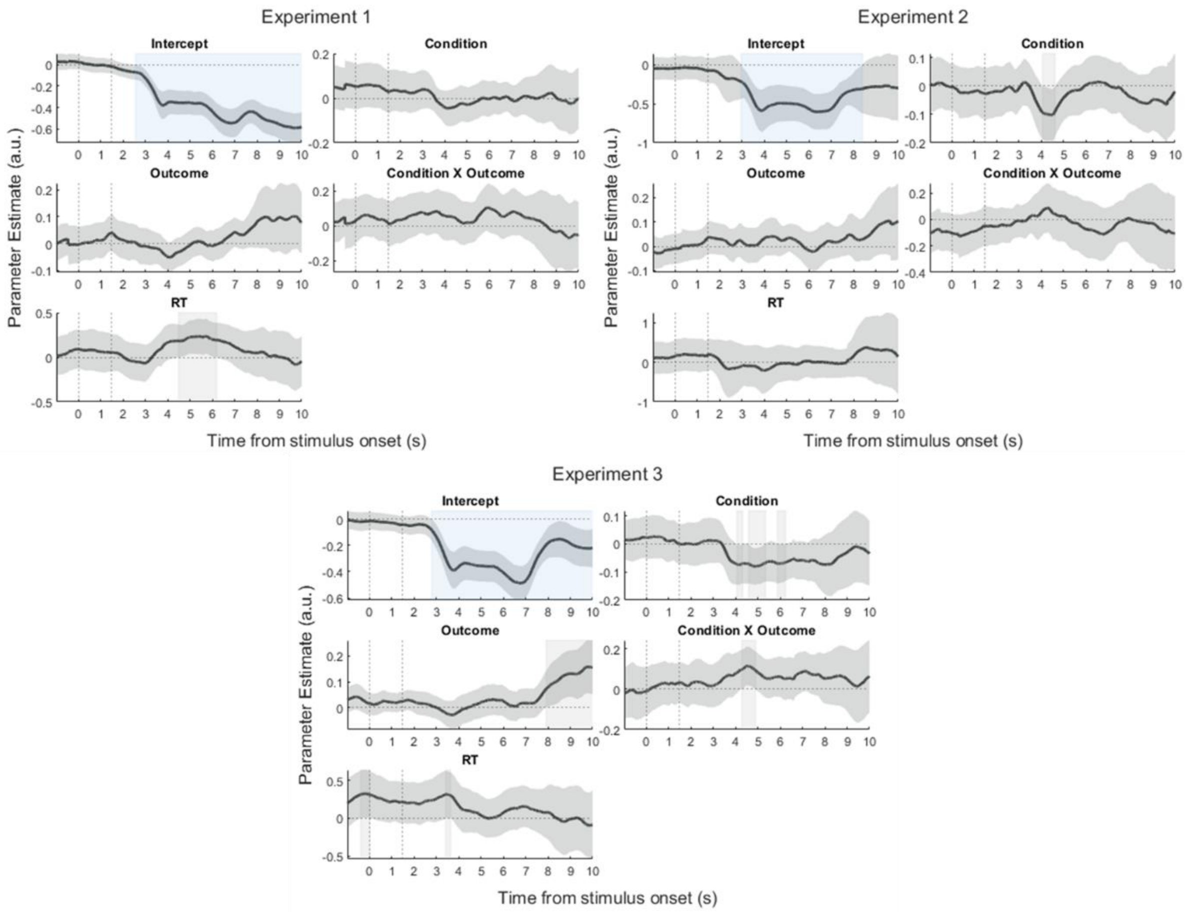

Figure F1: *GLM results on head rotation from the chosen face for each experiment. Positive values: looking away from the face of the avatar in the chosen elevator. Negative values: looking at the face of the avatar in the chosen elevator. Vertical dashed lines: doors' opening and end of the choice time. Blue and grey shaded areas: significant and non-significant clusters.*

G. Tables

Table G1: *Proportion of avoidance and response times for all experiments.*

		P(Avoid angry)		Approach RTs (ms)		Avoid RTs (ms)	
	condition	Prop	SD	median	MAD	median	MAD
Exp1 (n=60)	Unpred	.647	.478	557.133	330.738	752.044	309.882
	Pred	.726	.446	529.169	310.258	772.930	278.480
Exp2 (n=30)	Unpred	.599	.490	543.341	371.164	682.385	392.671
	Pred	.659	.474	515.276	330.376	766.066	350.532
Exp3 (n=60)	Unpred	.574	.495	473.414	268.142	598.837	371.420
	Pred	.613	.487	445.221	227.364	668.103	351.250

Note: Unpred and Pred: Unpredictable and Predictable conditions. Prop: Proportion of avoidance responses. SD: Standard Deviation. MAD: Median Absolute Dispersion.

Table G2: *Sample description and Self-reported measures across experiments.*

	Experiment 1 (n=60)		Experiment 2 (n=30)		Experiment 3 (n=60)	
	Stat	Cronbach α	Stat	Cronbach α	Stat	Cronbach α
Age	23 [18, 35]		22 [18, 33]		22 [18, 35]	
Gender	30F/30M		15F/15M		32F/28M	
Handedness	6L/53R/1A		3L/27R		4L/55R/1A	
Education	6A/11B/43C		4A/4B/22C		14A/5B/41C	
LSAS	37.0 [6, 91]	.91	32.5 [0, 63]	.90	44.5 [2, 85]	.94
STAI-S	28.0 [20, 50]	.89	32.0 [20, 51]	.89	29.0 [20, 47]	.81
STAI-T	42.0 [24, 73]	.90	41.5 [28, 64]	.87	43.0 [27, 64]	.84
BIS	21.0 [12, 26]	.68	20.5 [12, 28]	.80	20.0 [14, 26]	.60
BAS-DR	9.0 [5, 13]	.59	9.0 [6, 14]	.60	9.0 [5, 14]	.56
BAS-FS	12.0 [7, 16]	.55	13.0 [8, 16]	68	12.0 [7, 16]	.64
BAS-RR	17.0 [11, 19]	.55	17.0 [12, 19]	.51	17.0 [11, 20]	.52
PHQ-8	4.0 [0, 15]	.69	5.5 [0, 12]	.70	6.0 [0, 19]	.75
PANAS-P	33.0 [18, 44]	.88	33.5 [25, 42]	.70	33.0 [14, 43]	.77
PANAS-N	16.0 [10, 40]	.81	17.0 [10, 33]	.84	17.5 [10, 35]	.84
Cybersickness	7.0 [0, 26]	.81	6.0 [0, 30]	.90	8.5 [0, 25]	.76

Note: Stat: descriptive statistic, median [min, max] for continuous variables, count for categorical variables. Handedness: L: left-handed, R: right-handed, A: Ambidextrous. Education: A: high-school level education, B: two years of education after high-school, C: more than two years of education after high-school. LSAS: Liebowitz Social Anxiety Scale. STAI-S and -T: State and Trait subscales of the State-Trait Anxiety Inventory. BIS: Behavioral Inhibition Scale. BAS-DR, -FS and -RR: Behavioral Approach Scale, Drive, Fun Seeking and Reward Responsiveness *subscales. PHQ-8: Patient Health Questionnaire 8. PANAS-P and -N: Positive and Negative Affective Scale, Positive and Negative subscales. Cybersickness: simulator sickness questionnaire.*

H. Metanalytic analyses

To test the reliability of DDM parameters of interest across the three experiments, we conducted a Bayesian random effects meta-analysis on our results using the R-package brms (Bürkner, 2017). We opted for a Bayesian approach for meta-analysis as it has been shown to outperform frequentist ones when a limited number of studies is available (Reis et al., 2023). Bayesian models have been fitted through 4 chains of 30000 iterations each (15000 used as warm up period) and a thinning factor of 10. We used a *Normal(0,1)* prior for the mu parameter and a *HalfCauchy(0, 0.5)* prior for the standard deviation parameter. The analyses have been run on parameters' estimates from the three experiments and the relative measurement error computed as (upper limit - lower limit of parameters' estimates 95%CrI) / 3.92. To be able to compare the three experiments, for Experiment 2 we used parameters from a DDM fitting the Condition effect on drift rate, boundary separation and non-decision time, even though this was not the best fitting model in this experiment.

When testing the drift rates, we found credible meta-analytic effects both for the intercept (i.e., drift rate in the Unpredictable condition: estimate=.408; 95%CrI=[.013, .738]; $p_{<0}$= .023; R-hat=1.001; ESS=5055) and for the effect of Predictability (estimate=.275; 95%CrI=[.009, .545]; $p_{<0}$= .024; R-hat=1.001; ESS=5053). Conversely, we did not find credible metanalytic effects for the effect of Condition neither on the boundary separation (estimate=.027; 95%CrI=[-.095, .136]; $p_{<0}$=.177; R-hat=1; ESS=4656) nor on the non-decision time (estimate=.001; 95%CrI=[-.069, .076]; $p_{<0}$= .485; R-hat=1; ESS=3757).

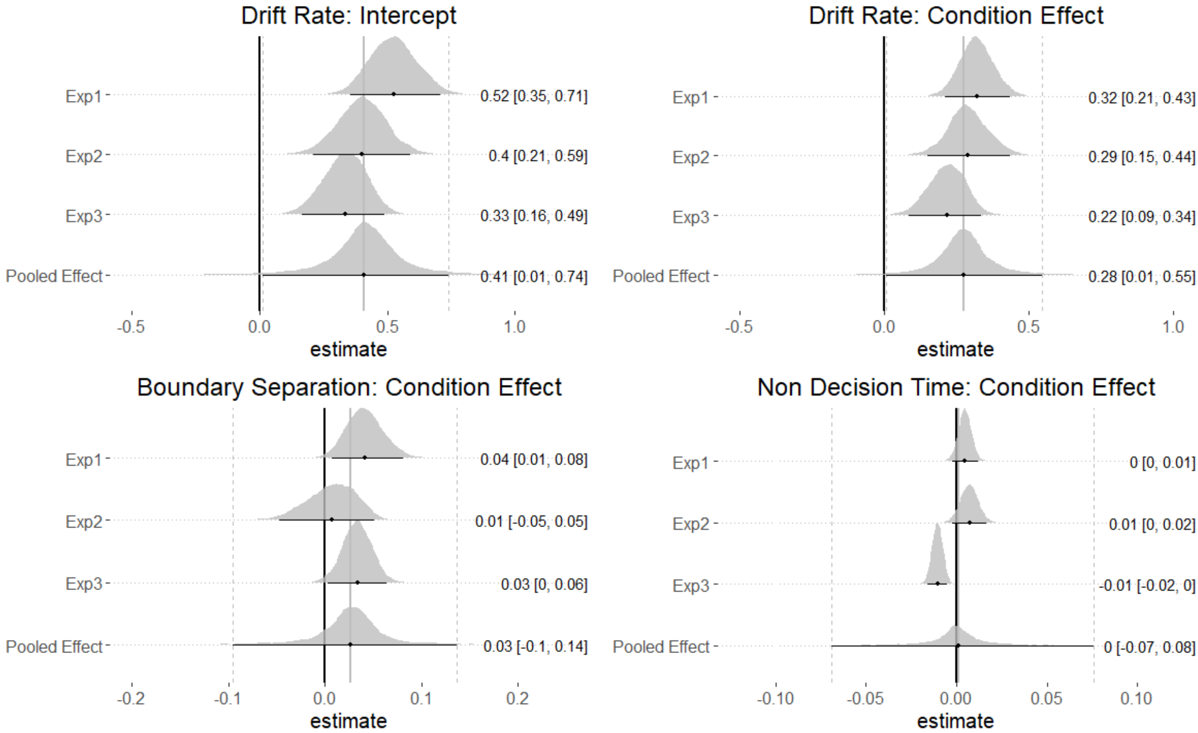

Figure H1: *Forest plots for each meta-analytic effect. Density distributions: posterior densities for parameter estimates across the three experiments and pooled effect. Vertical thin lines: median (solid line) and 95%CrI (dashed lines) of pooled effect distribution. Numbers on the right: posterior median estimate and 95%CrI (in brackets) for each parameter. Note that the parameters reported in the figure for the three experiments are the shrinked parameters after hierarchical meta-analytical estimation.*

I. Supplementary Figures

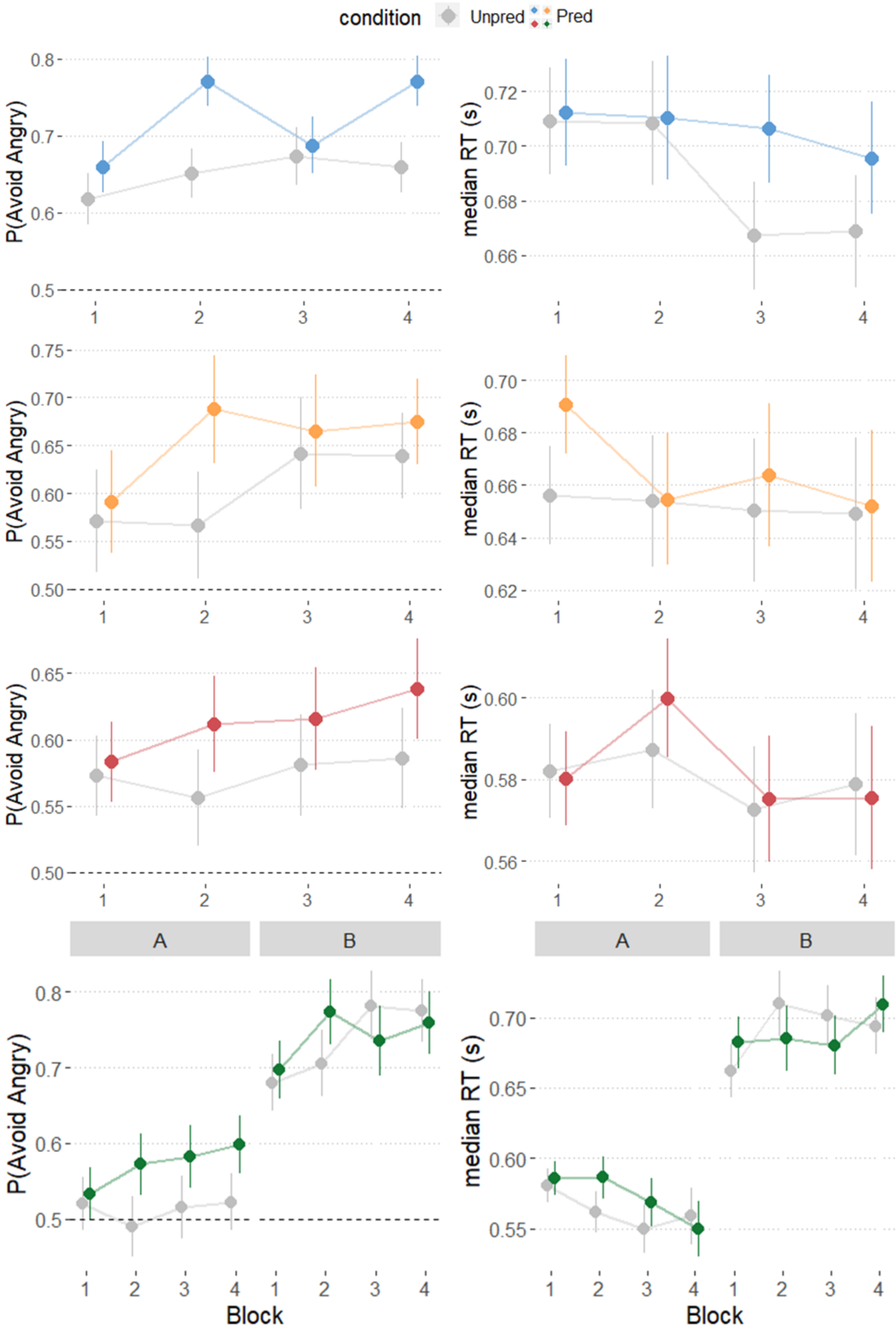

Figure I1: *Proportion of avoidance responses (left) and median response times (right) by Condition (Unpredictable vs Predictable) for each experiment and by class (Experiment 2 and 3 combined). Vertical lines: within subjects 95% confidence interval.*

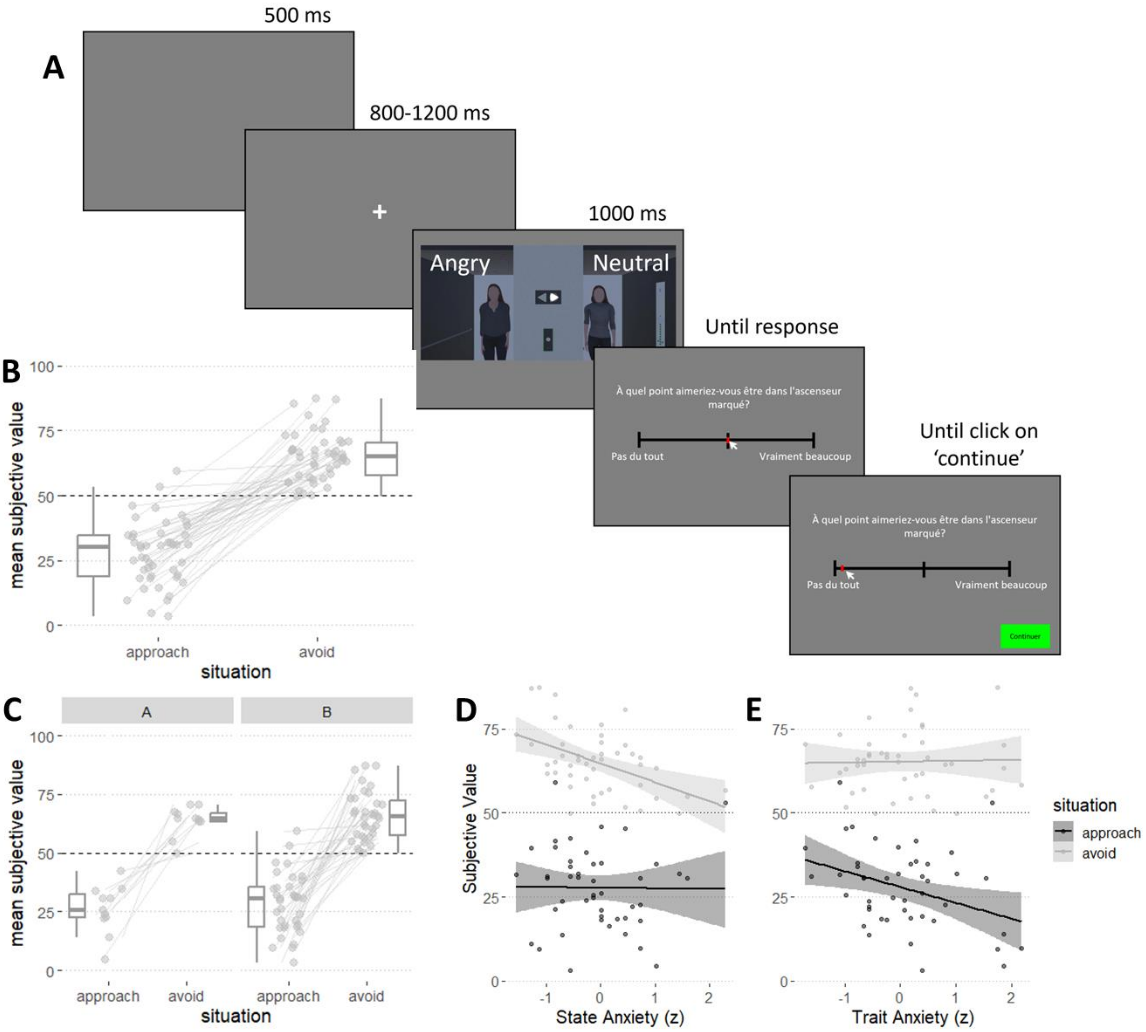

Figure I2: *(A) In the subjective evaluation task that followed the VR task, participants had to subjectively evaluate each possible choice scenario seen during the main task (the example is of an avoidance scenario). Note that the "Angry" and "Neutral" texts are for representational purpose only and were not present in the task. (B) Participants consistently preferred to avoid the angry avatar compared to approaching, with no difference between the two latent classes (C). (D) Individual reported subjective value for avoiding the angry individuals but not for approaching them correlated with state anxiety self-reported measures at the STAI; while subjective value for approaching but not for avoiding correlated with trait anxiety (E).*

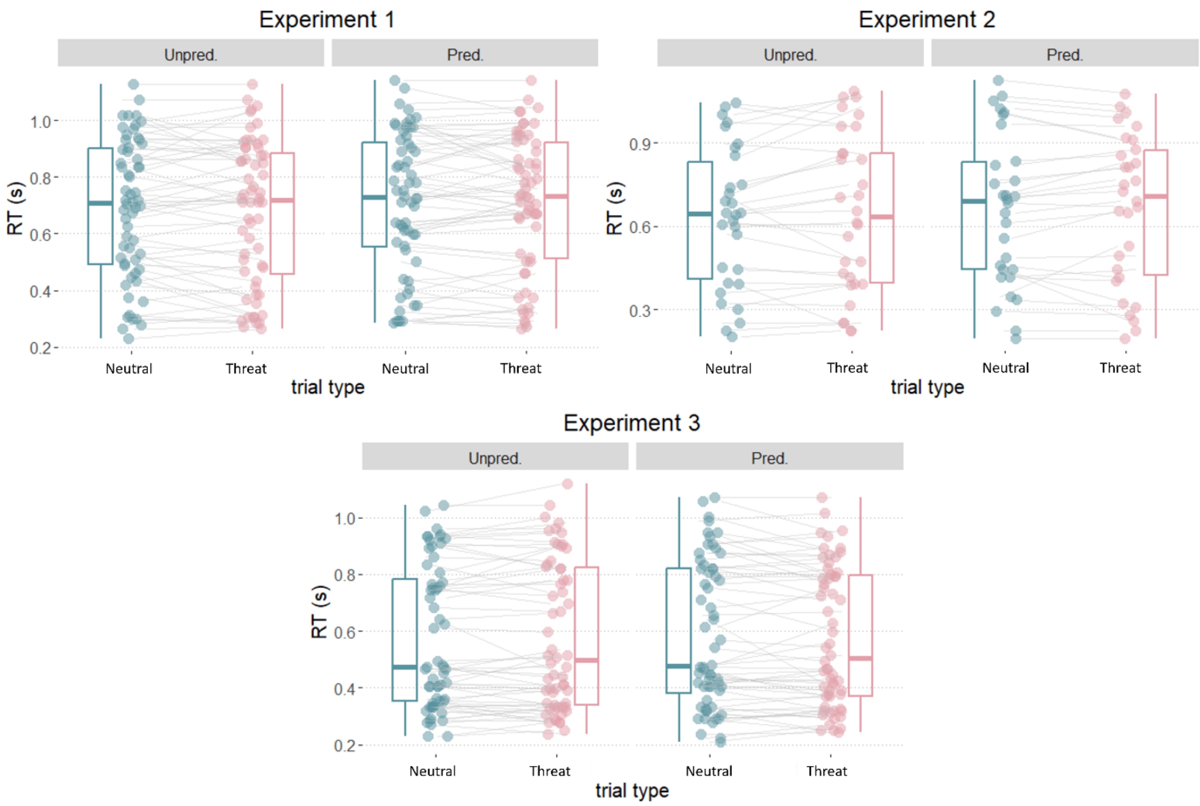

Figure I3: *Median response times by Condition (Unpredictable vs Predictable) and trial type (Neutral vs Threat) for each experiment.*

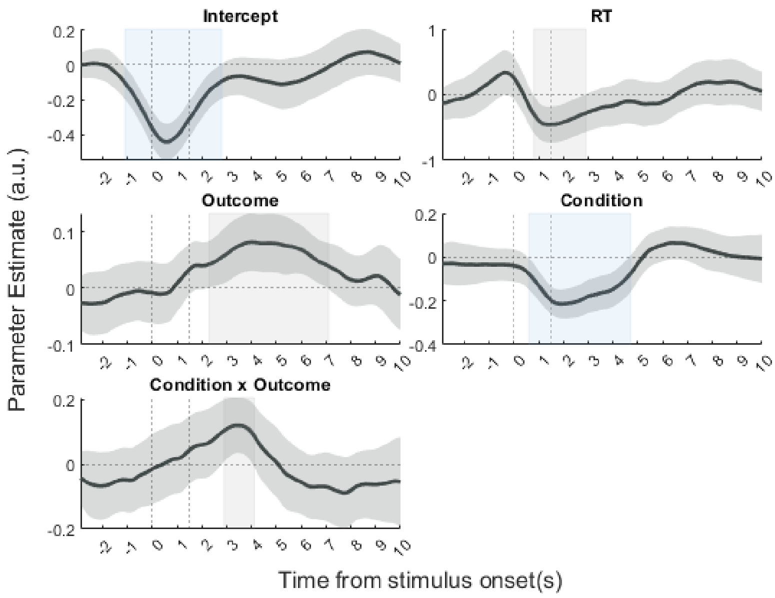

Figure I4: *GLM results on iHR in Experiment 3. Blue shaded areas are significant clusters, grey areas are non-significant clusters.*

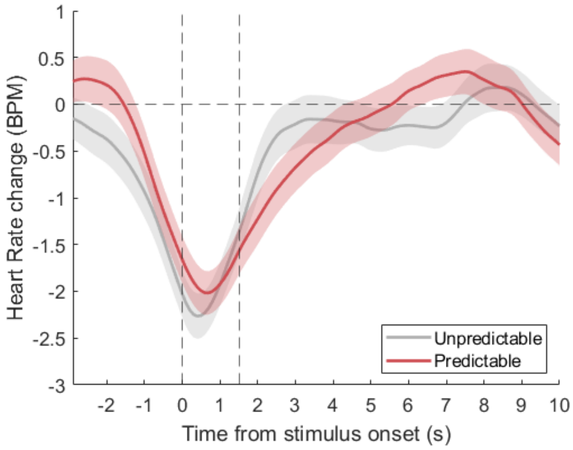

Figure I5: *Grand Average iHR series (baseline corrected) by Condition in neutral trials of Experiment 3. Vertical Lines: Stimulus onset (i.e. doors' opening) and end of choice time. Shaded areas: 95% within subject confidence intervals.*

Supplementary References

Brookes, J., Warburton, M., Alghadier, M., Mon-Williams, M., & Mushtaq, F. (2020). Studying human behavior with virtual reality: The Unity Experiment Framework. *Behavior Research Methods*, *52*(2), 455–463. https://doi.org/10.3758/s13428-019-01242-0

Bürkner, P.-C. (2017). brms: An *R* Package for Bayesian Multilevel Models Using *Stan*. *Journal of Statistical Software*, *80*(1). https://doi.org/10.18637/jss.v080.i01

Cignoni, P., Corsini, M., & Ranzuglia, G. (2008). *MeshLab: An Open-Source 3D Mesh Processing System*.

Feng, Y., Feng, H., Black, M. J., & Bolkart, T. (2021). *Learning an Animatable Detailed 3D Face Model from In-The-Wild Images* (arXiv:2012.04012). arXiv. http://arxiv.org/abs/2012.04012

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion*, *24*(8), 1377–1388. https://doi.org/10.1080/02699930903485076

Matzke, D., & Wagenmakers, E.-J. (2009). Psychological interpretation of the ex-Gaussian and shifted Wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review*, *16*, 798–817. https://doi.org/10.3758/PBR.16.5.798

Treal, T., Jackson, P. L., Jeuvrey, J., Vignais, N., & Meugnot, A. (2021). Natural human postural oscillations enhance the empathic response to a facial pain expression in a virtual character. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-91710-5

Wiecki, T., Sofer, I., & Frank, M. (2013). HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Frontiers in Neuroinformatics*, *7*. https://www.frontiersin.org/articles/10.3389/fninf.2013.00014