Distributional Learning is Associated with Modified Risk Preferences

Supplementary materials

Table of Contents: page

Supplementary Tables

Supplementary Table 1: Sensitivity to mean outcome, variance and skewness in reported estimates.

Results of single-sample t-tests assessing whether sensitivity is significantly greater than 0 across estimated distribution and estimated worth**.**

Study	Task	Measure	Parameter	Mean	SD	t(df)	p value	Cohen's d
Initial	SO	Distribution	Mean	1.065	0.192	t(153)= 68.9	< .001	5.55
Replication	SO	Distribution	Mean	1.075	0.206	t(210)= 75.9	< .001	5.226
Initial	SO	Distribution	Variance	0.834	0.361	t(153)= 28.7	< .001	2.312
Replication	SO	Distribution	Variance	0.798	0.332	t(210)= 35	< .001	2.406
Initial	SO	Distribution	Skewness	1.158	0.664	t(153)= 21.6	< .001	1.743
Replication	SO	Distribution	Skewness	1.205	0.748	t(210)= 23.4	< .001	1.611
Initial	SO	Worth	Mean	1.04	0.349	t(153)= 37	< .001	2.983
Replication	SO	Worth	Mean	1.062	0.388	t(210)= 39.8	< .001	2.739
Initial	SO	Worth	Variance	0.8	0.671	t(153)= 14.8	< .001	1.191
Replication	SO	Worth	Variance	0.819	0.663	t(210)= 17.9	< .001	1.235
Initial	SO	Worth	Skewness	0.757	0.78	t(153)= 12.1	< .001	0.971
Replication	SO	Worth	Skewness	0.756	0.736	t(210)= 14.9	< .001	1.028
Initial	Bandit	Distribution	Mean	0.684	0.417	t(153)= 20.4	< .001	1.641
Replication	Bandit	Distribution	Mean	0.674	0.473	t(210)= 20.7	< .001	1.425
Initial	Bandit	Distribution	Variance	0.246	0.393	t(153)= 7.8	< .001	0.626
Replication	Bandit	Distribution	Variance	0.326	0.484	t(210)= 9.8	< .001	0.673
Initial	Bandit	Distribution	Skewness	0.185	0.484	t(153)= 4.7	< .001	0.382
Replication	Bandit	Distribution	Skewness	0.241	0.544	t(210)= 6.4	< .001	0.443
Initial	Bandit	Worth	Mean	0.692	0.501	t(153)= 17.1	< .001	1.38
Replication	Bandit	Worth	Mean	0.698	0.508	t(210)= 20	< .001	1.374
Initial	Bandit	Worth	Variance	0.243	0.525	t(153)= 5.8	< .001	0.464
Replication	Bandit	Worth	Variance	0.38	0.583	t(210)= 9.5	< .001	0.651
Initial	Bandit	Worth	Skewness	0.053	0.477	t(153)= 1.4	0.172	0.11
Replication	Bandit	Worth	Skewness	0.162	0.499	t(210)= 4.7	< .001	0.324

Supplementary Table 2: Risk preferences and estimation biases.

Results of single-sample t-tests assessing whether estimation biases (for estimated distribution and estimated worth) and preferences (for choices and predicted happiness ratings) for high versus low variance are significantly different than 0**.**

Study	Task	Measure	Distribution mean	Bias	SD	t(df)	p value	Cohen's d
Initial	SO	Distribution	Low	-0.1	0.862	t(81)=-2.6	0.011	-0.11
Initial	SO	Distribution	High	0.03	0.907	t(77)=1.1	0.296	0.03
Replication	SO	Distribution	Low	-0.02	0.895	t(101)=-0.6	0.567	-0.02
Replication	SO	Distribution	High	0.05	0.983	t(108)=1.7	0.093	0.05
Initial	SO	Worth	Low	-0.05	1.138	t(81)=-1.7	0.096	-0.05
Initial	SO	Worth	High	0.07	0.622	t(77)=2.6	0.011	0.12
Replication	SO	Worth	Low	0	1.174	t(101)=-0.1	0.963	0
Replication	SO	Worth	High	0.08	0.729	t(108)=2.4	0.017	0.1
Initial	SO	Choices	Low	0.24	1.002	t(81)=2.7	0.009	0.24
Initial	SO	Choices	High	-0.27	0.935	t(77)=-3.1	0.002	-0.29
Replication	SO	Choices	Low	0.04	1.035	t(101)=0.5	0.635	0.04
Replication	SO	Choices	High	-0.19	1.019	t(108)=-2.4	0.019	-0.18
Initial	SO	Happiness	Low	0.14	0.923	t(66)=4.5	< .001	0.15
Initial	SO	Happiness	High	-0.12	0.828	t(70)=-4.9	< .001	-0.15
Replication	SO	Happiness	Low	0.13	0.85	t(98)=4.8	< .001	0.15
Replication	SO	Happiness	High	-0.13	1.036	t(106)=-4.1	< .001	-0.12
Initial	Bandit	Distribution	Low	-0.05	0.987	t(77)=-1.3	0.194	-0.05
Initial	Bandit	Distribution	High	0.13	0.842	t(81)=4.2	< .001	0.16
Replication	Bandit	Distribution	Low	0.01	1.012	t(108)=0.3	0.755	0.01
Replication	Bandit	Distribution	High	0.06	0.919	t(101)=2.2	0.031	0.06
Initial	Bandit	Choices	Low	-0.24	1.029	t(74)=-3.7	< .001	-0.24
Initial	Bandit	Choices	High	0.22	0.854	t(77)=4	< .001	0.25
Replication	Bandit	Choices	Low	-0.1	1.008	t(102)=-1.8	0.073	-0.1
Replication	Bandit	Choices	High	0.19	0.966	t(95)=3.4	< .001	0.2
Initial	Bandit	Happiness	Low	-0.04	0.98	t(70)=-0.8	0.421	-0.04
Initial	Bandit	Happiness	High	0.07	0.92	t(66)=1.8	0.073	0.08

Replication	Bandit	Happiness	Low	-0.01	1.072	t(106)=-0.3	0.758	-0.01
Replication	Bandit	Happiness	High	-0.04	1.095	t(98)=-1.5	0.148	-0.04
Initial	Bandit	Worth	Low	-0.04	1.067	t(77)=-0.9	0.376	-0.03
Initial	Bandit	Worth	High	0.1	0.846	t(81)=3.2	0.002	0.12
Replication	Bandit	Worth	Low	0.04	1.102	t(108)=1.3	0.197	0.03
Replication	Bandit	Worth	High	0.04	0.894	t(101)=1.4	0.168	0.04

Supplementary Table 3: Differences in sensitivity between tasks.

Results of paired-sample t-tests comparing the difference in sensitivity between tasks.

Study	Measure	Parameter	Mean	SD	t(df)	p value	Cohen's d
Initial	Distribution	expectancy	0.381	0.421	t(153)= 11.2	< .001	0.906
Initial	Worth	expectancy	0.348	0.571	t(153)= 7.6	< .001	0.61
Initial	Distribution	variance	0.588	0.541	t(153)= 13.5	< .001	1.086
Initial	Worth	variance	0.556	0.789	t(153)= 8.7	< .001	0.704
Initial	Distribution	skewness	0.973	0.76	t(153)= 15.9	< .001	1.281
Initial	Worth	skewness	0.704	0.852	t(153)= 10.3	< .001	0.827
Replication	Distribution	expectancy	0.401	0.514	t(210)= 11.3	< .001	0.78
Replication	Worth	expectancy	0.364	0.508	t(210)= 10.4	< .001	0.716
Replication	Distribution	variance	0.473	0.543	t(210)= 12.6	< .001	0.87
Replication	Worth	variance	0.439	0.724	t(210)= 8.8	< .001	0.606
Replication	Distribution	skewness	0.964	0.763	t(210)= 18.3	< .001	1.263
Replication	Worth	skewness	0.595	0.83	t(210)= 10.4	< .001	0.716

Supplementary Tables 4: Differences in sensitivity between tasks, controlling for task order

Results of linear regression models predicting within-subject differences in sensitivity between tasks, using **task order** as a predictor. The significance of the intercept reflects the effect of sensitivity differences between tasks, controlling for task order.

<u>Sensitivity to expectancy:</u>

Estimated distribution	Initial study			
Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.339	0.033	-10.222	< .001
Phase order – bandit task first	-0.093	0.066	-1.411	0.16

Estimated distribution	Replication			
Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.323	0.052	-6.171	< .001
Phase order – bandit task first	-0.141	0.071	-2.002	0.047

Estimated worth	Initial study			
Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.295	0.046	-6.44	< .001
Phase order – bandit task first	-0.142	0.091	-1.568	0.12

Estimated worth	Replication			
Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.333	0.052	-6.384	< .001
Phase order – bandit task first	-0.056	0.07	-0.798	0.43

Sensitivity to variance:

Estimated distribution	Initial study			
Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.553	0.043	-12.92	< .001
Phase order – bandit task first	-0.063	0.085	-0.744	0.46

Estimated distribution	Replication			
Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.424	0.056	-7.621	< .001
Phase order – bandit task first	-0.088	0.075	-1.168	0.24

Estimated worth	Initial study			
Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.445	0.061	-7.285	< .001
Phase order – bandit task first	-0.265	0.121	-2.195	0.03

Estimated worth	Replication			
Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.413	0.074	-5.542	< .001
Phase order – bandit task first	-0.048	0.1	-0.481	0.63

<u>Sensitivity to skewness:</u>

Estimated distribution	Initial study			
Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.758	0.056	-13.551	< .001
Phase order – bandit task first	-0.491	0.111	-4.44	< .001

Estimated distribution	Replication			
Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-1.09	0.078	-14.046	< .001
Phase order – bandit task first	0.229	0.105	2.19	0.03

Estimated worth	Initial study			
Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.644	0.068	-9.514	< .001
Phase order – bandit task first	-0.178	0.134	-1.326	0.19

Estimated worth	Replication			
Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.664	0.085	-7.803	< .001
Phase order – bandit task first	0.126	0.115	1.101	0.27

Supplementary Tables 5: Differences in sensitivity between tasks, controlling for temporal distance between observation and report

Results of linear regression models predicting within-subject differences in sensitivity between tasks, using the **sum of chest indices** (standardized) as a predictor. Each sensitivity measure is based on the differences in reported parameters between chest pairs (high vs. low variance or positive vs. negative skewness). If forgetfulness due to the time between observation and report reduces sensitivity, we would expect sensitivity to decline as the sum of the chest pair indices increases. The significance of the intercept reflects the predicted effect of sensitivity differences between tasks, controlling for the impact of forgetfulness.

Sensitivity to variance:

Estimated distribution	Initial study			
Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.588	0.043	-13.538	<0.001
Chest indices sum	-0.054	0.034	-1.58	0.116

Estimated distribution	Replication			
Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.473	0.037	-12.66	<0.001
Chest indices sum	-0.038	0.029	-1.327	0.186

Estimated worth	Initial study			
Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.556	0.064	-8.735	<0.001
Chest indices sum	-0.043	0.05	-0.853	0.395

Estimated worth	Replication			
Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.439	0.05	-8.798	<0.001
Chest indices sum	-0.029	0.039	-0.758	0.449

Sensitivity to skewness:

Estimated distribution	Initial study			
Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.973	0.061	-15.9	<0.001
Chest indices sum	-0.047	0.048	-0.976	0.331

Estimated distribution	Replication			
Predictor	Estimate (B)	Std. Error (SE)	t value	p value

Intercept	-0.964	0.052	-18.443	<0.001
Chest indices sum	0.072	0.04	1.776	0.077

Estimated worth	Initial study			
Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.704	0.068	-10.336	<0.001
Chest indices sum	0.097	0.054	1.805	0.073

Estimated worth	Replication			
Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.595	0.057	-10.398	<0.001
Chest indices sum	-0.036	0.044	-0.824	0.411

Supplementary Table 6: Sensitivity to mean outcome in preference measures

Results of single-sample t-tests assessing whether preference measures were significantly different for higher mean-outcome chests. Choice preferences were measured as the percentage of high mean-outcome chests selected and tested against chance level (0.5). Predicted happiness preferences were defined as the mean difference in predicted happiness ratings between high and low mean-outcome chests.

Study	Task	Measure	Preference	SD	t(df)	p value	Cohen's d
Initial	SO	Choices	0.78	0.22	t(159)=16.2	< .001	1.27
Initial	SO	Happiness	0.62	0.28	t(159)=27.9	< .001	2.21
Initial	Bandit	Choices	0.82	0.14	t(159)=28.4	< .001	2.29
Initial	Bandit	Happiness	0.4	0.3	t(159)=17.2	< .001	1.33
Replication	SO	Choices	0.73	0.2	t(210)=17	< .001	1.15
Replication	SO	Happiness	0.65	0.27	t(210)=34.6	< .001	2.41
Replication	Bandit	Choices	0.79	0.14	t(210)=30.3	< .001	2
Replication	Bandit	Happiness	0.41	0.28	t(210)=21	< .001	1.46

Supplementary Tables 7: Differences between estimation bias and risk preferences in the Sequential Observation (SO) task

The analysis aims to assess whether the effect of chest mean outcome differs between estimation biases and preferences. This could be represented using the formula:

$$(est_{\mu_+} - est_{\mu_-}) - (pref_{\mu_+} - pref_{\mu_-})$$

Where est and $pref$ represent the estimation biases and preferences for variance, respectively, and μ_- and μ_+ represent cases where the chests have a positive or negative mean.

This can be rearranged as:

$$(est_{\mu_+} - pref_{\mu_+}) - (est_{\mu_-} - pref_{\mu_-})$$

This approach enables us to calculate differences within each mean outcome condition before making between-condition comparisons. Since each participant experienced only one mean outcome condition in the SO task, this method allows for a within-subject computation of differences. We then use a between-subjects analysis, specifically – multiple linear regression controlling for the effects of task order, to determine whether this within-subject measure varies across mean outcome conditions, which is a between-subjects variable.

The results of the multiple regressions, for each measure comparison (estimates vs subjective values) in each study are reported bellow:

Estimated distribution: Happiness Initial study

Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	0.592	0.1	5.933	< .001
Chest mean outcome – positive	-1.148	0.122	-9.403	< .001
Task order – bandit first	0.084	0.138	0.611	0.54

Estimated distribution: Happiness Replication

Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	0.585	0.11	5.329	< .001
Chest mean outcome – positive	-0.866	0.125	-6.933	< .001
Task order – bandit first	-0.247	0.125	-1.967	0.05

Estimated distribution: Choices Initial study

Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.29	0.102	-2.849	0.005
Chest mean outcome – positive	0.702	0.129	5.435	< .001
Task order – bandit first	-0.243	0.148	-1.641	0.102

Estimated distribution: Choices Replication

Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	0.23	0.12	1.919	0.056
Chest mean outcome – positive	-0.377	0.136	-2.765	0.006
Task order – bandit first	-0.065	0.137	-0.474	0.636

Estimated worth : Happiness Initial study

Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.608	0.103	-5.88	< .001
Chest mean outcome – positive	1.084	0.127	8.565	< .001
Task order – bandit first	0.106	0.143	0.741	0.460

Estimated worth : Happiness Replication

Predictor	Estimate (B)	Std. Error (SE)	t value	p values
Intercept	0.498	0.113	4.411	< .001
Chest mean outcome – positive	-0.78	0.128	-6.075	< .001
Task order – bandit first	-0.169	0.129	-1.312	0.190

Estimated worth : Choices Initial study

Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.382	0.1	-3.806	< .001
Chest mean outcome – positive	0.806	0.127	6.336	< .001
Task order – bandit first	-0.089	0.146	-0.607	0.544

Estimated worth : Choices Replication

Predictor	Estimate (B)	Std. Error (SE)	t value	p values
Intercept	0.174	0.12	1.444	0.150
Chest mean outcome - positive	-0.346	0.137	-2.527	0.012
Task order – bandit first	0.008	0.137	0.061	0.951

Supplementary Tables 8: Task Differences in the differential impact of mean outcome on estimation biases and preferences

For each participant, we calculated the difference between estimation biases and preferences (both scaled) and summed these differences across both tasks. We then used multiple linear regression to predict this measure, with the mean of both normally distributed chests in the SO task as the predictor and task order as a covariate.

This analysis was based on the following rationale. Reformulating our measure from the previous analysis:

$$(\Delta_{\mu_+} - \Delta_{\mu_-}) = \left(est_{\mu_+} - \ pref_{\mu_+}\right) - \ \left(est_{\mu_-} - \ pref_{\mu_-}\right)$$

we define the current measure of interest as:

$$\left(SO_{\Delta_{\mu_+}} - SO_{\Delta_{\mu_-}}\right) - \ \left(bandit_{\Delta_{\mu_+}} - bandit_{\Delta_{\mu_-}}\right)$$

This captures task differences in how mean outcome differentially influences estimation biases and preferences.

Since each participant encountered opposite mean-outcome conditions in each task, we reordered the expression to match together opposing mean-outcome conditions across tasks:

$$\left(SO_{\Delta_{\mu_+}} + \ bandit_{\Delta_{\mu_-}}\right) - \ \left(bandit_{\Delta_{\mu_+}} + SO_{\Delta_{\mu_-}}\right)$$

This expression captures the difference between tasks by focusing on the effect of mean-outcome condition in the SO task (and correspondingly the opposite condition in the bandit task) on the sum of the differences between estimation biases and preferences for both tasks.

Estimated distribution: Happiness Initial study

Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	0.41	0.133	3.082	0.002
Chest mean in SO – positive	-0.818	0.158	-5.171	< .001
Task order – bandit first	0.032	0.165	0.191	0.849

Estimated distribution: Happiness Replication

Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	0.387	0.119	3.265	0.001
Chest mean in SO – positive	-0.496	0.135	-3.681	< .001
Task order – bandit first	-0.236	0.135	-1.74	0.083

Estimated distribution: Choices Initial study

Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.245	0.122	-2.002	0.047
Chest mean in SO – positive	0.643	0.151	4.266	< .001
Task order – bandit first	-0.203	0.159	-1.275	0.204

Estimated distribution: Choices Replication

Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	0.272	0.119	2.275	0.024
Chest mean in SO – positive	-0.563	0.137	-4.097	< .001
Task order – bandit first	0.037	0.138	0.269	0.788

Estimated worth : Happiness Initial study

Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.47	0.135	-3.49	< .001
Chest mean in SO – positive	0.776	0.16	4.838	< .001
Task order – bandit first	0.2	0.167	1.196	0.234

Estimated worth : Happiness Replication

Predictor	Estimate (B)	Std. Error (SE)	t value	p values
Intercept	0.343	0.118	2.907	0.004
Chest mean in SO – positive	-0.58	0.134	-4.322	< .001
Task order – bandit first	-0.076	0.135	-0.565	0.573

Estimated worth : Choices	Initial study			
Predictor	Estimate (B)	Std. Error (SE)	t value	p value
Intercept	-0.385	0.12	-3.21	0.002
Chest mean in SO – positive	0.782	0.148	5.288	< .001
Task order – bandit first	0.012	0.156	0.076	0.939

Estimated worth : Choices	Replication			
Predictor	Estimate (B)	Std. Error (SE)	t value	p values
Intercept	0.257	0.118	2.182	0.03
Chest mean in SO – positive	-0.646	0.135	-4.771	< .001
Task order – bandit first	0.145	0.136	1.067	0.287

Computational models

Supplementary Table 9: additional model comparisons for the SO task.

For the first model comparison (A), we evaluated two additional variants to rule out the influence of specific modeling assumptions. The first variant removed attention bias by setting $\lambda = 0$. The second variant altered the learning process: instead of learning full outcome distributions, the models maintained two separate point estimates, one for the expected value (Q) and the other for expected utility (Q') which were updated over time using the following equations:

(6) $$Q_{t+1} = Q_t + \eta(V_t - Q_t)$$
(7) $$Q'_{t+1} = Q'_t + \eta(u(V_t) - Q'_t)$$

with $u(V_t)$ representing utility-transformed value.

Integrated Bayesian inference criteria (iBIC) values for additional models of the SO task

Study	Model set	A_1	A_{0a}	A_{0b}
Initial	No attention bias	3235	3371	3409
Initial	Point estimates	3234	3343	3405
Initial	Standard	3143	3304	3369
Replication	No attention bias	5121	5278	5346
Replication	Point estimates	5030	5127	5175
Replication	Standard	4962	5051	5111

Supplementary Table 10: additional model comparison for the differences between tasks.

For the second model comparison (B), we tested a different model family to ensure our results did not hinge on the assumption of a constant learning rate in the bandit task. These models implemented a decaying learning rate by weighting each new outcome as an exponential function of its absolute value, resembling an equally weighted average over time. The update rule was:

(8) $$Q_{t+1} = \frac{Q_t * W_t + V_t * e^{\lambda |V_i|}}{W_t + e^{\lambda |V_i|}}$$

where W_t is the cumulative weight of past outcomes:

(9) $$W_{t+1} = W_t + e^{\lambda |V_i|}$$

The expected utility estimate (Q') in model B_0 followed the same structure, but with utility-transformed values:

(10) $$Q'_{t+1} = \frac{Q'_t * W_t + u(V_t) * e^{\lambda |V_i|}}{W_t + e^{\lambda |V_i|}}$$

As stated in the main text, we also evaluated variants of model B_0 and B_1 where learning bias was allowed to differ between the SO and bandit tasks (models B_2 and B_3 respectively).

Integrated Bayesian inference criteria (iBIC) values for additional models of the differences between tasks

Study	Model set	B_0	B_1	B_2	B_3
Initial	Full information	2825	2779	2820	2777
Initial	Standard	2736	2649	2677	2640
Replication	Full information	5008	4959	4900	4850
Replication	Standard	4932	4820	4824	4754

Supplementary Table 11: Skew preferences and estimation biases.

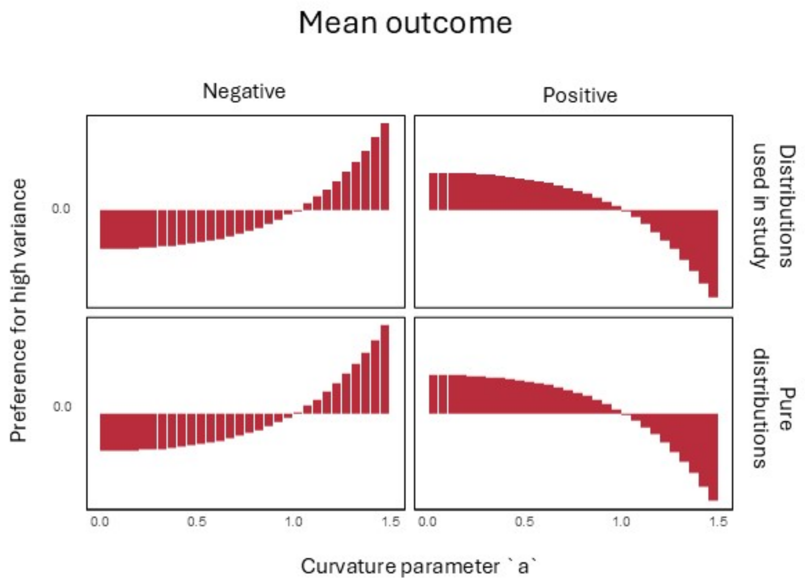
Results of single-sample t-tests assessing whether estimation biases (for estimated distribution and estimated worth) and preferences (for choices and predicted happiness ratings) in favor of positive versus negative skew are significantly different than 0**.**

Study	Task	Measure	Distribution mean	Bias	SD	t(df)	p value	Cohen's d
Initial	SO	Distribution	Low	0.62	1	t(81)=5.6	<.001	0.11
Initial	SO	Distribution	High	0.73	0.95	t(77)=6.8	<.001	0.22
Replication	SO	Distribution	Low	0.45	0.59	t(101)=7.6	<.001	0.38
Replication	SO	Distribution	High	0.61	1.07	t(108)=6	<.001	0.1
Initial	SO	Worth	Low	-0.1	1	t(81)=-0.9	0.353	-0.02
Initial	SO	Worth	High	0.06	0.86	t(77)=0.6	0.555	0.03
Replication	SO	Worth	Low	-0.06	0.7	t(101)=-0.9	0.38	-0.05
Replication	SO	Worth	High	0.06	1.11	t(108)=0.6	0.562	0.01
Initial	SO	Choices	Low	0.25	0.99	t(81)=2.3	0.022	0.21
Initial	SO	Choices	High	0.04	1.06	t(77)=0.3	0.732	0.03
Replication	SO	Choices	Low	0.3	1.05	t(101)=2.9	0.005	0.23
Replication	SO	Choices	High	0.36	0.99	t(108)=3.8	<.001	0.29
Initial	SO	Happiness	Low	0.46	1.08	t(66)=3.5	<.001	0.13
Initial	SO	Happiness	High	0.56	0.91	t(70)=5.1	<.001	0.16
Replication	SO	Happiness	Low	0.74	1.03	t(98)=7.2	<.001	0.16
Replication	SO	Happiness	High	0.53	0.82	t(106)=6.7	<.001	0.22
Initial	Bandit	Distribution	Low	0.24	0.85	t(77)=2.5	0.014	0.1
Initial	Bandit	Distribution	High	0.07	1.06	t(81)=0.6	0.527	0.02
Replication	Bandit	Distribution	Low	-0.03	0.98	t(108)=-0.4	0.718	-0.01
Replication	Bandit	Distribution	High	0.01	0.94	t(101)=0.1	0.95	0
Initial	Bandit	Worth	Low	0.16	0.73	t(77)=1.9	0.063	0.08
Initial	Bandit	Worth	High	0.11	0.07	t(81)=0.9	0.362	0.03
Replication	Bandit	Worth	Low	-0.06	0.71	t(108)=-0.9	0.366	-0.02
Replication	Bandit	Worth	High	-0.07	1.06	t(101)=-0.6	0.531	-0.02
Initial	Bandit	Choices	Low	-0.13	1	t(74)=-1.2	0.252	-0.07
Initial	Bandit	Choices	High	-0.3	1.05	t(77)=-2.5	0.015	-0.15
Replication	Bandit	Choices	Low	-0.2	0.97	t(102)=-2	0.044	-0.1
Replication	Bandit	Choices	High	-0.34	1.07	t(95)=-3.1	0.003	-0.16
Initial	Bandit	Happiness	Low	0.09	0.86	t(70)=0.9	0.393	0.03
Initial	Bandit	Happiness	High	-0.05	1	t(66)=-0.4	0.699	-0.01
Replication	Bandit	Happiness	Low	0.07	1.1	t(106)=0.6	0.531	0.02
Replication	Bandit	Happiness	High	-0.07	1.03	t(98)=-0.7	0.513	-0.02

Supplementary Figures

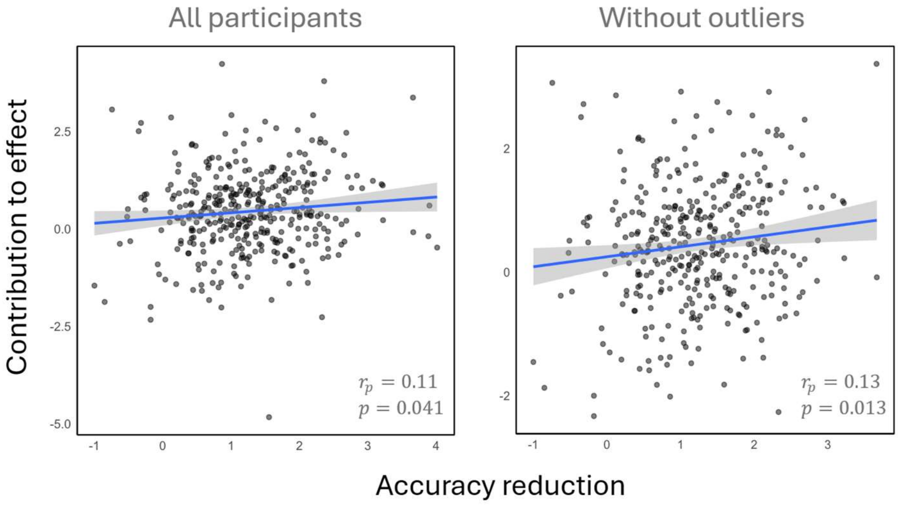
Supplementary Figure 1: Predicted risk preferences.

Predicted preferences for high- vs. low-variance outcome distributions across different values of the utility function curvature parameter `$a$` (as defined in Eq. 2). Predictions are presented separately by mean outcome condition (positive vs. negative) and by distribution type: *symmetric distributions* used in the study, which included both positive and negative outcomes, and *pure distributions*, in which all outcomes fell entirely within the domain of gains or losses. For pure distributions, high-variance outcomes ranged from −10 to 0 (negative) and 0 to 10 (positive), while low-variance outcomes ranged from −6 to −4 (negative) and 4 to 6 (positive).



As shown in Figure 5 in the main article, the curvature parameters that best fit the data fell within the range [0, 1].

Supplementary Figure 2: Correlation between accuracy reduction and contribution to the effect

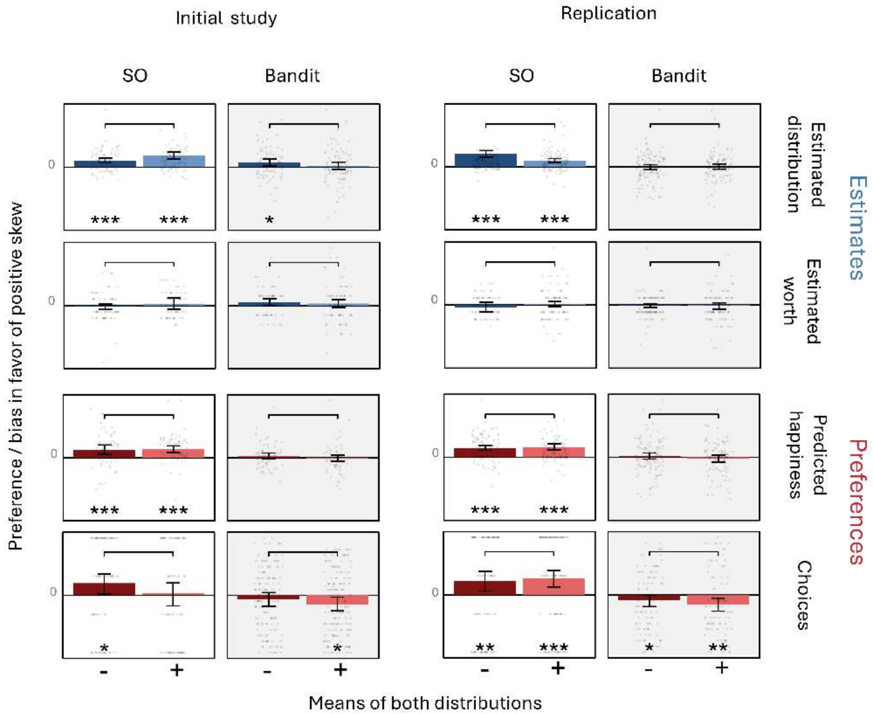


A scatterplot illustrating the relationship between participants' reduction in accuracy from the SO task to the bandit task and their contribution to task-related differences in estimation bias and preference pattern divergence. Accuracy reduction was calculated as the difference in accuracy between tasks, measured using both estimated distribution (Hellinger's distance) and estimated worth (RMSE). These two measures were standardized (scaled to have a variance of 1 across all participants) and then averaged to create a single accuracy reduction score per participant.

Contribution to the effect was determined by computing the mean differences in the influence of mean outcome across pairs of estimation and preference measures. This value was summed across both tasks and reversed for participants who experienced a negative mean outcome for the symmetric chests in the SO task (and consequently, a positive mean outcome in the bandit task), following the formula outlined in Supplementary Table 8.

Pearson correlation was calculated both with and without outliers in either measure (Z-scores with an absolute value greater than 3) to ensure the correlation was not driven by specific data points.

Supplementary Figure 3: Preference patterns between skewed outcome distribution chests and corresponding estimation biases



The figure shows mean estimation biases and preferences favoring positive, vs negative, skewness across different measures. The data were split by task and by whether the two distributions' mean was positive or negative. Data were z-scored for comparability across measures. Error bars represent 95% CIs. Points represent individual participants. Asterisks below columns denote significant difference from 0 (*: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$), whereas brackets indicate the insignificance of difference between columns.

Supplementary Notes

Supplementary Note 1: Skewness preference and estimation bias results

As described in the main article, in the SO task, participants favored positively skewed over negatively skewed chests, as commonly observed in description-based decision tasks, whereas in the bandit task, these preferences were eliminated (in predicted happiness measure) or even reversed (in participants' choices), as commonly observed in experience-based decision tasks (Supplementary Figure 3; Supplementary Table 11).

If preferences in favor of positively skewed outcomes in the SO task stemmed from overweighting of low probability outcomes, as posited by prospect theory, it is to be expected that participants' estimates would show similar biases. In this respect, the results were ambiguous. Whereas the reported distributions showed higher estimates for positively skewed, as compared to negatively skewed, distributions, this effect was absent in participants' estimated worths. Thus, we cannot rule out that the effect observed in the reported distribution is not due to an estimation error, but rather to a bias in the reporting of individual outcome probabilities. In this latter case, skewness preferences may be explained similarly to risk preferences, that is, as a result of the application of the utility function described in Eq. 2. This equation, when applied with negative `a` values, as was found to best fit the data in the computational modeling, induces a preference for positively skewed outcome distributions regardless of the distribution's mean outcome.

The difference between predicted happiness and choices in the bandit task, whereby only the latter favored negatively skewed chests, is consistent with the common explanation of experience-based choices as reflecting a recency bias. Although the average expected value at any time during the task should be equal for the two chests, the overweighting of recent outcomes means that the expected value of the negatively skewed chests is more frequently greater than that of the positively skewed chests. This could have led participants to choose the former chests more frequently. The same effect was not expected, and indeed was not observed, in the predicted happiness measure because this measure is continuous, and thus its average should reflect the chests' average expected values. As noted, recency bias is typical of non-distributional learning algorithms that use a fixed learning rate to update expected outcomes. Furthermore, even happiness rating–based preferences, which were not expected to be biased by a recency bias, did not exhibit the positive skewness preference that would be expected if the utility function were effectively applied. This supports the claim the effective application of a utility function depends on distributional learning.

Thus, while somewhat ambiguous, these results appear to support two parallel effects of compromised distributional learning: the underweighting of rare events and the ineffective application of participants' utility function.