PNAS
www.pnas.org

Supplementary Information for:

Pairing Facts with Imagined Consequences Improves Pandemic-Related Risk Perception

Alyssa H. Sinclair[1,2]*, Shabnam Hakimi[1]*, Matthew Stanley[1,2], R. Alison Adcock[1,2,3], & Gregory R. Samanez-Larkin[1,2]

[1]Duke University, Center for Cognitive Neuroscience, [2]Duke University, Department of Psychology and Neuroscience, [3]Duke University, Department of Psychiatry and Behavioral Sciences

Corresponding Author: Alyssa H. Sinclair, ahs50@duke.edu

This PDF file includes:

Supplementary Text

Figures S1 to S7

Tables S1 to S6

Other supplementary materials for this manuscript include the following:

Data and code provided here: https://osf.io/35us2/

Supplementary Information: Appendix

Statistical Analysis

All analyses were conducted in R (v4.0), implemented in RStudio (v1.3.1093). Figures were generated using the sjPlot (1) and ggplot2 (2) packages.

In Study 1, we calculated Pearson product-moment correlations to assess relationships between pairs of variables. The measure of actual risk (prevalence-based exposure risk) was log-transformed to normalize the distribution and meet assumptions for parametric statistical tests.

In Study 2, continuous variables were standardized before submission to multiple linear regression. Factor variables for conditions were effect coded. Visual inspection of histograms indicated that several variables exhibited high kurtosis, with some extreme values at both tails of the distribution. As a result, residuals from fitted models were larger for values at the tails. To correct for high kurtosis and meet the assumption of normality, we winsorized extreme values to the 5th and 95th percentiles. Variables for change in perceived risk (Session 1) and change in willingness to engage in risky activities (Session 1 and Session 2) were winsorized. Winsorizing these variables improved model fits but did not change the statistical significance of any of our findings (Supplemental Material, *Results without Winsorizing*). Additionally, we corrected skewed distributions by applying log-transformations to the variables for actual risk, retrospective report of risky behaviors, and willingness to engage in risky activities (Session 1). Other variables were not transformed because distributions were approximately normal.

Exclusions

We excluded all data from 88 participants for the following preregistered reasons: lack of COVID-19 statistics for their location (27 participants), failing an attention check (27 participants), or providing irrelevant or excessively short responses to the Episodic Simulation task (34 participants). We also excluded two extreme outlier observations for the retrospective report of risky behaviors between sessions (15/15 activities) because it was exceptionally unlikely that any participant could have completed the full list of activities over the course of a week (e.g., going to the dentist, getting a haircut, and flying on an airplane). Manual inspection of the data from these participants indicated that their other responses appeared legitimate, suggesting that they may have misread the instructions for this particular question. Therefore, we omitted their responses for this question, but did not exclude other data from these participants. Lastly, 35 participants failed to complete all questions for the Risk Estimation task during the Session 2 follow-up survey. These incomplete data points were excluded from the analysis of risk estimation accuracy.

Everyday Activities Assessed

In Study 1, we measured **perceived risk** by asking participants to rate the subjective perceived riskiness of engaging in six different everyday activities in their local communities, using a 5-point Likert scale (*1 = Not at all risky … 5 = Extremely risky*). The activities were as follows: Going for a walk outside, shopping at a grocery store, eating inside a restaurant, meeting with a small group of friends, travelling within one's state, or travelling beyond one's state. We measured **willingness to engage in risky activities** by asking participants to report (yes/no) whether they would be willing to participate in eight different activities, if hypothetically all stay-at-home restrictions in their area were lifted. The activities were as follows: Going to a park or playground, going to the gym, eating inside a restaurant, meeting with up to 5 friends, meeting with up to 10 friends, meeting with over 10 friends, travelling within one's state, or travelling beyond one's state.

In Study 2, we measured **perceived risk** by asking participants to rate the subjective perceived riskiness of engaging in 15 different everyday activities in their local communities, using a 5-point Likert scale (*1 = Not at all risky … 5 = Extremely risky*). We also measured **willingness** by asking participants to rate their willingness to engage in each of these 15 activities, using a 5-point Likert scale (*1 = Definitely would NOT do this … 5 = Definitely WOULD do this*). The activities were as follows: Picking up takeout food, walking outside without a mask in an area without many people, having an outdoor picnic with friends 6+ feet apart, playing a group sport outside without a mask, grocery shopping indoors with a mask, retail shopping indoors with a mask, going to the dentist, taking a taxi/Uber/Lyft, dining outdoors at a restaurant, dining indoors at a restaurant, getting a haircut, exercising at a gym without a mask, flying on an airplane[1], going to an indoor bar or nightclub, or going to a large indoor house party.

Deviations from Preregistration

In addition to the demographic and individual difference measures that we tested as covariates (see above, *Controlling for Individual Differences*), we also measured Actively Open-Minded Thinking and Social Value Orientation. These measures were listed under planned exploratory analyses, but we did not analyze these individual differences here for the sake of brevity. As described above, in Study 2 we excluded two data points from the variable for retrospectively reported risky activities between testing sessions. All other exclusion criteria were preregistered. Furthermore, we included a covariate for delay length (between testing sessions) in all Session 2 models. This covariate was not preregistered because we did not

[1] Note that flying on an airplane may involve close contact with people from one's local community (e.g., fellow passengers), but could also include people from surrounding counties and other cities (e.g., in the airport). In case this ambiguity influenced our results, we also reported alternative results with this item omitted from the perceived risk scale (refer to subheading *Alternative Measure of Perceived Risk*).

anticipate receiving survey responses over a long period of time. We also computed a model-free measure that resolved global changes in risk misestimation across the full set of group sizes, area between the curves. This single metric provided a more parsimonious alternative to testing change in misestimation separately for each group size (as planned).

Equivalence Testing

We conducted Two One-Sided Tests of Equivalence, using the TOSTER package in R (3), to assess whether the association between perceived and actual risk was statistically equivalent to zero. Following current best practices (3), we set the upper and lower bounds for the equivalence tests to the smallest effect size that can be reliably detected (80% power), given the sample size. Study 1 and Study 2 analyses used different bounds because of the discrepancy in sample sizes.

In Study 1, the Pearson correlation between baseline perceived risk and actual risk was r = 0.047, with a sample size of 234. The bounds of the equivalence test were set to +/- 0.19. The equivalence test was significant ($p < .001$), whereas the null hypothesis test was non-significant ($p = .201$). Next, we performed the same test for willingness to engage in risky activities. The Pearson correlation between baseline willingness to engage in risky activities and actual risk was r = -0.012. The equivalence test was significant ($p < .001$), whereas the null hypothesis test was non-significant ($p = .744$). Taken together, these two pairs of tests indicate that the observed effects were statistically not different from zero and were statistically equivalent to zero.

In Study 2, the Pearson correlation between baseline perceived risk and actual risk was r = -0.00387, with a sample size of 735. The bounds of the equivalence test were set to +/- 0.11. The equivalence test was significant ($p = .002$), whereas the null hypothesis test was non-significant ($p = .917$). Next, we performed the same test for willingness to engage in risky activities. The Pearson correlation between baseline willingness to engage in risky activities and actual risk was r = -0.012. The equivalence test was significant ($p = .049$), whereas the null hypothesis test was non-significant ($p = .183$). Taken together, these two pairs of tests indicate that the observed effects were statistically not different from zero and were statistically equivalent to zero.

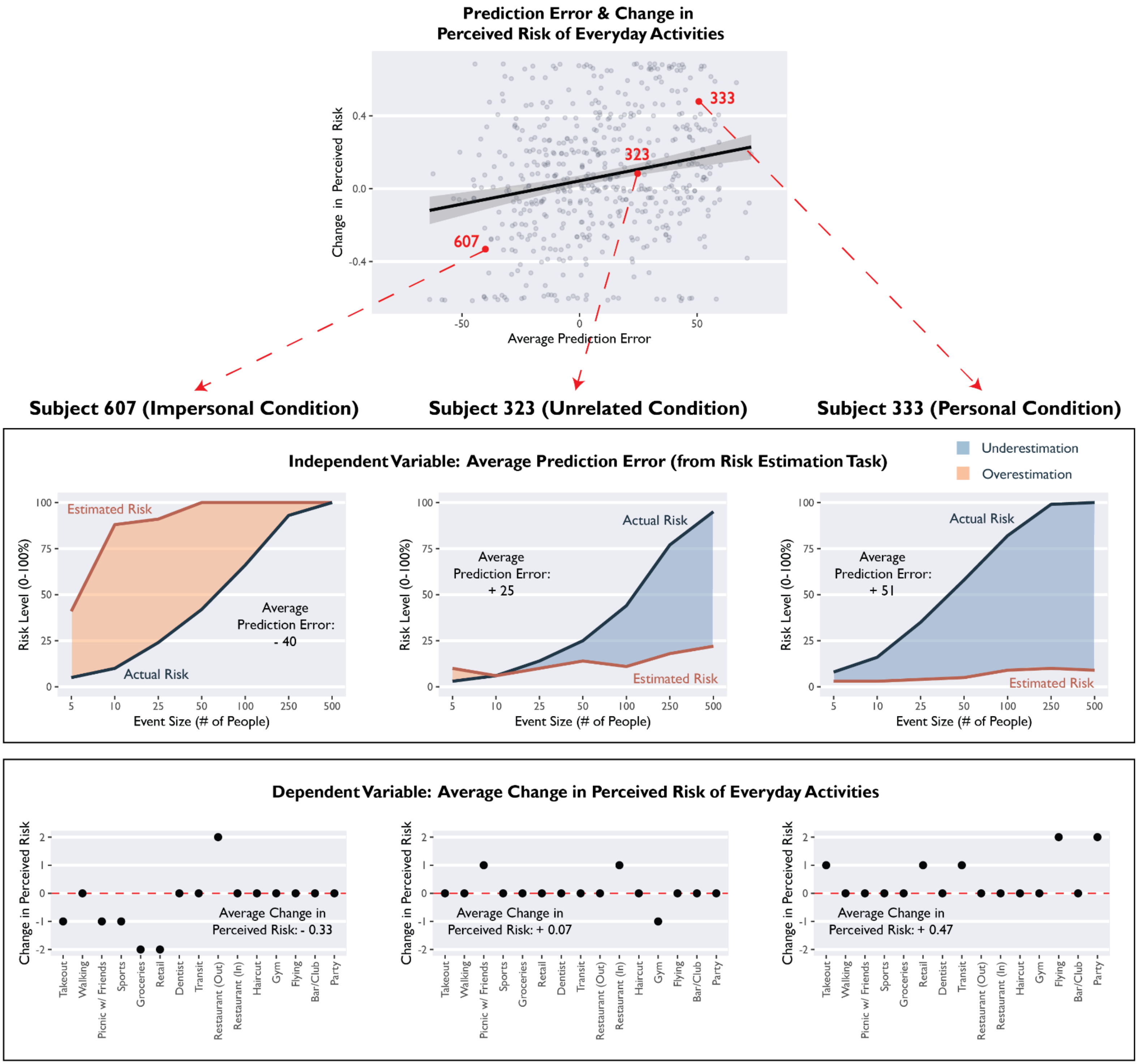

Figure S1. Breakdown of composite measures, visualizing all data from three example participants. Top: A reproduction of Figure 4A, with three example subjects highlighted in red. Middle: Curves depicting the discrepancy between actual risk and estimated risk, from the Risk Estimation task. Orange shaded areas indicate risk overestimation, and blue shaded areas indicate risk underestimation. Average prediction error scores were computed by averaging the misestimation bias (actual – estimated risk) across the seven event sizes. The average prediction error score derived from this analysis produces the x-coordinate of each point in the top scatterplot. Bottom: Changes in perceived risk (post-intervention – baseline) for each of the 15 everyday activities assessed. Average change in perceived risk was computed as the average change across the 15 items. The average change in perceived risk score produces the y-coordinate of each point in the top scatterplot.

Responders and Non-Responders

As described in the main text, we found that only a very small proportion of participants demonstrated a backfire effect in response to the intervention. We also identified participants who showed beneficial change after the intervention, based on their risk estimation bias (i.e., sign of the average prediction error score) at baseline. We classified participants as "Responders" if they were Underestimators who reported greater perceived risk and lower willingness to engage in risky activities after the intervention, or risk Overestimators who reported lower perceived risk and greater willingness. The slight majority of participants showed beneficial change after the intervention: 52.2% of participants (285/546) reported more accurate risk perception and 56.8% of participants (310/546) reported willingness to engage in risky activities that was better aligned with actual risk levels. We could not determine the prevalence of beneficial and backfire effects for the Unguided Exploration condition because these participants did not complete the Risk Estimation task; therefore, we could not classify these participants as Underestimators or Overestimators. Overall, we found that the efficacy of our interventions was not undermined by a backfire effect; the vast majority of participants either benefitted from the intervention or showed no change.

Why was the intervention ineffective for some participants? We hypothesized that a subset of non-responders may have already held accurate beliefs about risk at baseline, rendering the intervention unnecessary. To test this idea, we conducted an exploratory analysis comparing the average unsigned prediction error scores between responders (participants who shifted perceived risk in the appropriate direction) and non-responders (participants who did not). We found that non-responders were slightly more accurate than responders at baseline; non-responders reported lower unsigned prediction error scores, $t(535) = -2.05$, $p = .041$, Cohen's $d = -0.18$, 95% CI [-0.35, -0.01]. This effect was driven by a subset of non-responders (29.4%, 74/252) who were already highly accurate at risk estimation (average unsigned prediction error scores <= 15). Distributions of unsigned prediction error scores for responders and non-responders are provided in Figure S1.

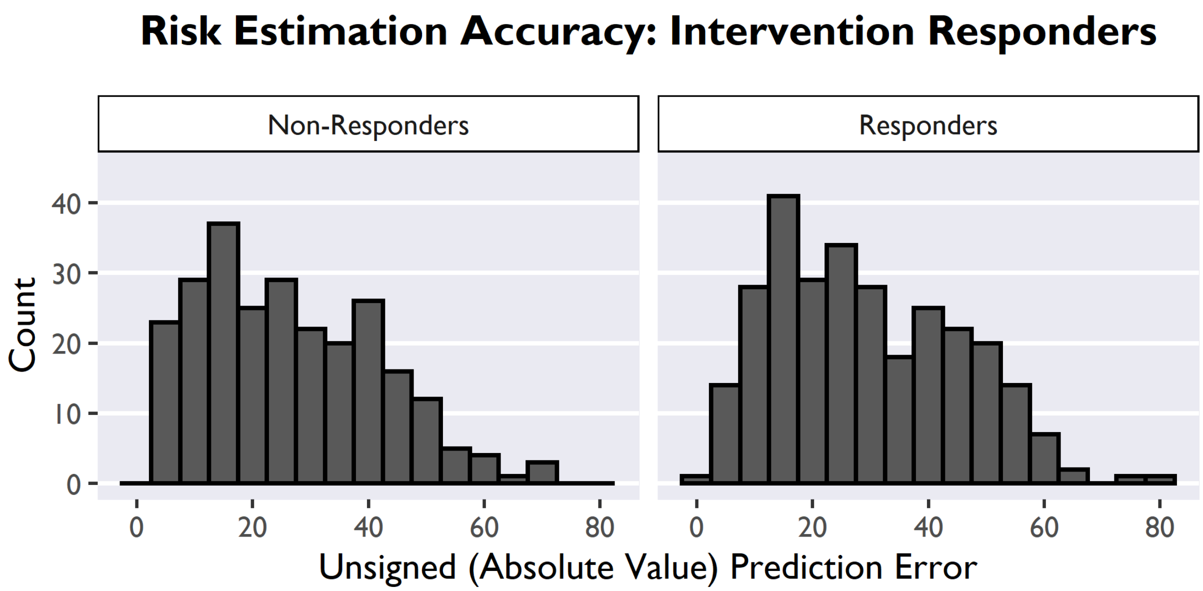

Figure S2. Distributions of average unsigned prediction error scores. A) Among participants who did not respond to the intervention (i.e., did not shift perceived risk in the correct direction), a subset showed very high risk estimation accuracy before feedback (prediction error scores <= 10), suggesting that the intervention was unnecessary. B) Among participants who did respond well to the intervention (i.e., shifted perceived risk in the correct direction), fewer participants showed very high baseline risk estimation accuracy.

Table S1: Summary Statistics

	Risk Underestimators (Average Prediction Error >=15)		
	Impersonal (N = 72)	**Personal (N = 77)**	**Unrelated (N = 77)**
Perceived Risk			
Baseline	2.90 ± 0.78	2.88 ± 0.98	2.86 ± 0.82
S1	3.08 ± 0.81	3.06 ± 0.94	2.94 ± 0.79
S2	3.09 ± 0.84	3.07 ± 0.92	2.96 ± 0.83
S1-Baseline Change	0.18 ± 0.34	0.18 ± 0.45	0.08 ± 0.38
S2-Baseline Change	0.19 ± 0.38	0.21 ± 0.39	0.09 ± 0.35
Willingness			
Baseline	2.87 ± 0.88	2.82 ± 0.88	2.90 ± 0.90
S1	2.58 ± 0.83	2.57 ± 0.90	2.71 ± 0.89
S2	2.72 ± 0.82	2.65 ± 0.80	2.79 ± 0.88
S1-Baseline Change	-0.29 ± 0.35	-0.25 ± 0.35	-0.19 ± 0.37
S2-Baseline Change	-0.15 ± 0.51	-0.15 ± 0.28	-0.08 ± 0.39
Prediction Error			
S1	37.13 ± 12.92	37.42 ± 14.59	38.07 ± 13.25

S2	6.60 ± 22.57	10.89 ± 23.71	14.33 ± 22.78

	Accurate Risk Estimators (Average Prediction Error -14—14)		
	Impersonal (N = 60)	**Personal (N = 59)**	**Unrelated (N = 64)**
Perceived Risk			
Baseline	3.09 ± 0.79	3.11 ± 0.72	3.27 ± 0.67
S1	3.17 ± 0.94	3.15 ± 0.70	3.36 ± 0.68
S2	3.16 ± 0.92	3.09 ± 0.75	3.36 ± 0.64
S1-Baseline Change	0.08 ± 0.43	0.04 ± 0.45	0.09 ± 0.33
S2-Baseline Change	0.06 ± 0.42	-0.03 ± 0.44	0.09 ± 0.39
Willingness			
Baseline	2.73 ± 0.91	2.75 ± 0.77	2.57 ± 0.65
S1	2.53 ± 0.95	2.64 ± 0.82	2.37 ± 0.62
S2	2.68 ± 0.91	2.66 ± 0.78	2.53 ± 0.71
S1-Baseline Change	-0.20 ± 0.28	-0.11 ± 0.31	-0.20 ± 0.26
S2-Baseline Change	-0.02 ± 0.33	-0.06 ± 0.35	-0.07 ± 0.45
Prediction Error			
S1	1.32 ± 7.89	-0.94 ± 8.48	-0.01 ± 9.07
S2	0.26 ± 19.00	-1.21 ± 18.04	-2.95 ± 18.30

	Risk Overestimators (Average Prediction Error <=15)		
	Impersonal (N = 45)	**Personal (N = 42)**	**Unrelated (N = 41)**
Perceived Risk			
Baseline	3.52 ± 0.60	3.34 ± 0.66	3.41 ± 0.76
S1	3.40 ± 0.66	3.34 ± 0.74	3.38 ± 0.77
S2	3.42 ± 0.67	3.32 ± 0.76	3.41 ± 0.72
S1-Baseline Change	-0.12 ± 0.44	-0.01 ± 0.40	-0.04 ± 0.35
S2-Baseline Change	-0.12 ± 0.53	0.05 ± 0.47	0.00 ± 0.35
Willingness			
Baseline	2.51 ± 0.68	2.51 ± 0.72	2.64 ± 0.83
S1	2.35 ± 0.71	2.37 ± 0.70	2.46 ± 0.68
S2	2.38 ± 0.59	2.47 ± 0.63	2.46 ± 0.65
S1-Baseline Change	-0.16 ± 0.35	-0.14 ± 0.30	-0.18 ± 0.54
S2-Baseline Change	-0.11 ± 0.38	-0.06 ± 0.33	-0.07 ± 0.34
Prediction Error			
S1	-30.19 ± 12.68	-29.20 ± 11.45	-27.92 ± 8.19
S2	4.36 ± 25.69	10.02 ± 20.33	10.74 ± 19.31

Table S1. Values indicate means ± standard deviations for key variables, subset by condition and misestimation type (Underestimators, Accurate Estimators, and Overestimators). Baseline = Pre-Intervention, S1 = Session 1 Post-Intervention, S2 = Session 2 Post-Intervention. S1-Baseline Change and S2-Baseline Change measures reflect the average within-subjects change.

Underestimators report increased perceived risk and decreased willingness to engage in risky activities, more so in the Impersonal and Personal conditions than in the Unrelated condition. Overestimators report decreases in perceived risk in the Impersonal condition, more so than in the Personal or Unrelated conditions. Accurate Estimators show slight increases in perceived risk. Refer to Figure S3 below for a visualization of these point estimates.

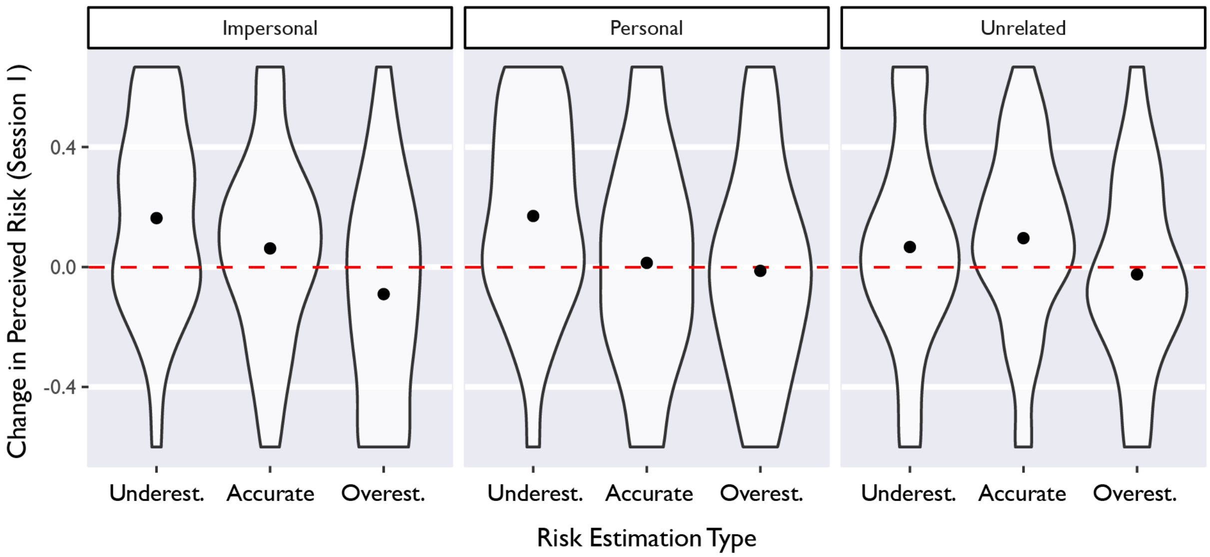

Figure S3. Violin plot depicting the average change in perceived risk (Session 1) across conditions, for participants who tended to underestimate, overestimate, or accurately estimate risk. This plot visualizes the descriptive statistics reported in the Table S1 above. Dots represent means, and areas around the dots represent the density of the distribution. Red dotted line indicates zero, no change from baseline.

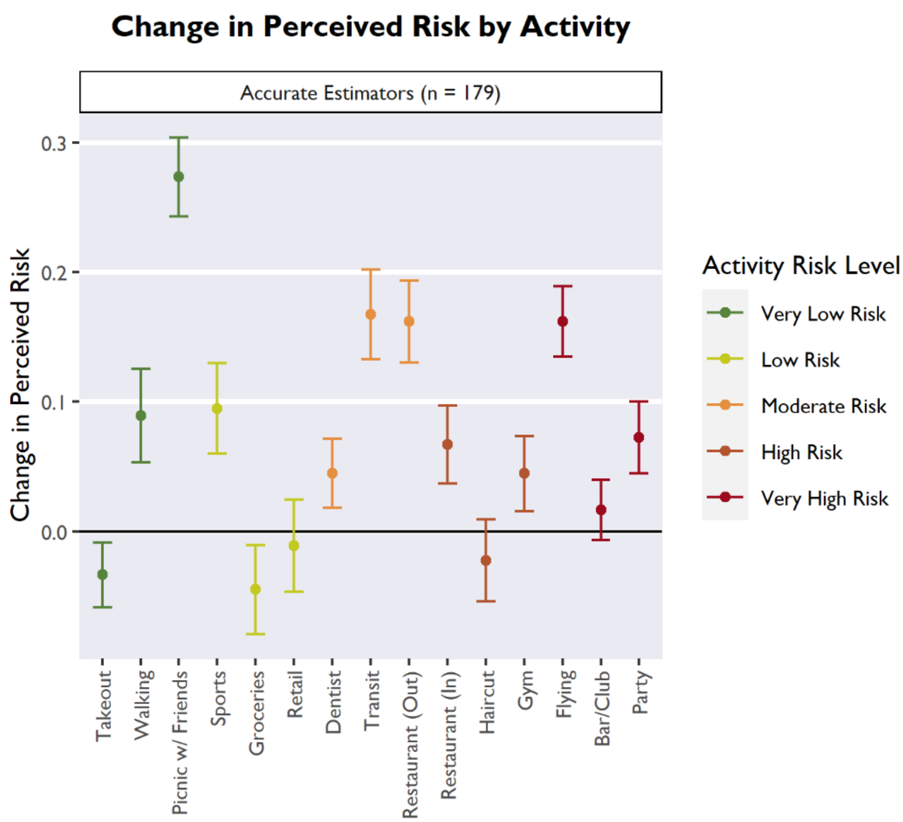

Figure S4. Average within-subjects change in perceived risk for each of the 15 everyday activities. This plot depicts the subset of participants who were relatively accurate at the risk estimation task (average prediction error score between -15 and +15). This plot can be compared with Figure 3. Colors indicate approximate risk levels of the activities (4). Error bars depict 95% confidence intervals around the means.

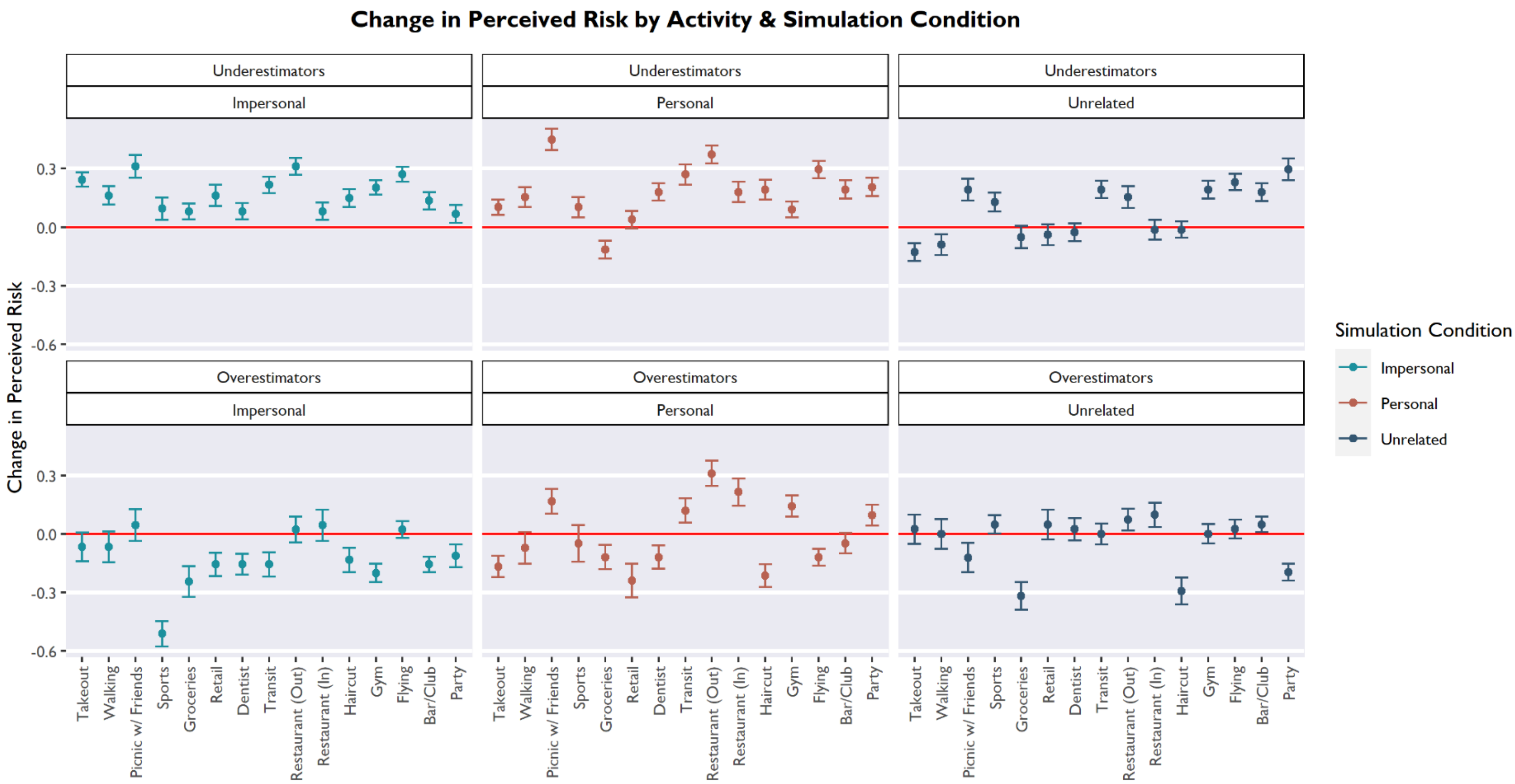

Figure S5. Average change in perceived risk for each of the 15 everyday activities assessed, split by Underestimator/Overestimator classification and Simulation Condition. Error bars indicate 95% confidence intervals. Red lines indicate zero (i.e., no change from baseline).

Regression to the Mean

Regression to the mean can give rise to illusory treatment effects when extreme values (due to noise or measurement error) tend to converge towards the population mean over time, due to chance alone. To rule out regression to the mean, we examined whether extreme baseline measurements of perceived risk were associated with greater change in perceived risk. Visual inspection of scatterplots revealed no evidence of a U-shaped function whereby extreme baseline measurements lead to greater change (Figure S6, S7).

We also conducted t-tests that assessed whether regression to the mean was evident. We binned participants by their baseline scores for perceived risk (on the 1-5 point scale): “Extreme” baseline values were defined as baseline perceived risk scores < 2 or > 4 ($n = 149$), whereas “Center” baseline values were defined as baseline perceived risk scores within the range [2, 4] ($n = 586$). If regression to the mean occurred, then participants with Extreme baseline scores would report greater change scores (absolute value). This would indicate that extreme values were more likely to move towards the mean. We found that in Session 1, there was no significant difference in the average change in perceived risk (absolute value) between Extreme and Center participants ($t(733) = -0.66$, $p = 0.506$, 95% CI [-0.07, 0.04]; Mean Center: 0.297, Mean

Extreme: 0.315). Likewise, in Session 2, there was no significant difference in the average lasting change in perceived risk (absolute value) between Extreme and Center participants ($t(669) = 0.97$, $p = 0.331$, 95% CI [-0.03, 0.08]; Mean Center: 0.337, Mean Extreme: 0.310). We also conducted the same analyses for each intervention condition separately, and did find any statistically significant differences between Extreme and Center means (all *p*s > .10 before correction for multiple comparisons). Overall, these null results demonstrate that regression to the mean cannot account for our intervention effects.

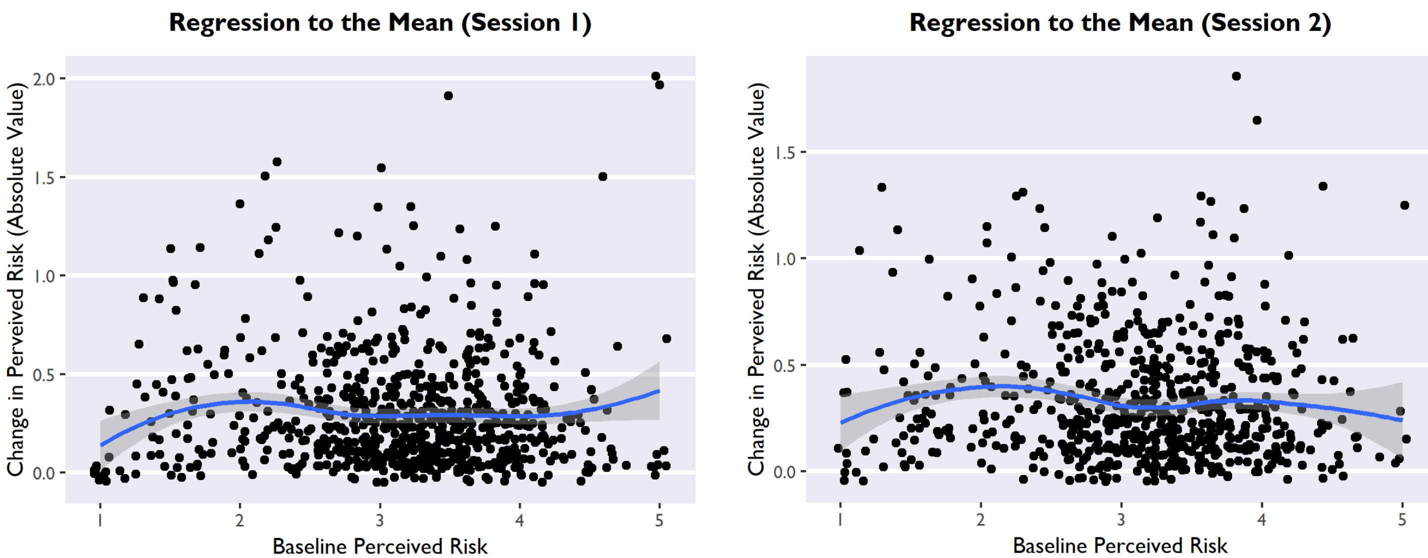

Figure S6. Scatterplot demonstrating the association between each participant's baseline perceived risk score and their post-intervention change in perceived risk, subset by session. If regression to the mean is present, then participants who reported extreme values of perceived risk at baseline (x-coordinates < 2 or > 4) should demonstrate more change (greater y-values). This pattern would produce a U-shaped function. These plots do not reveal a U-shaped function, demonstrating that participants tended to report similar amounts of change in perceived risk, regardless of their baseline measures.

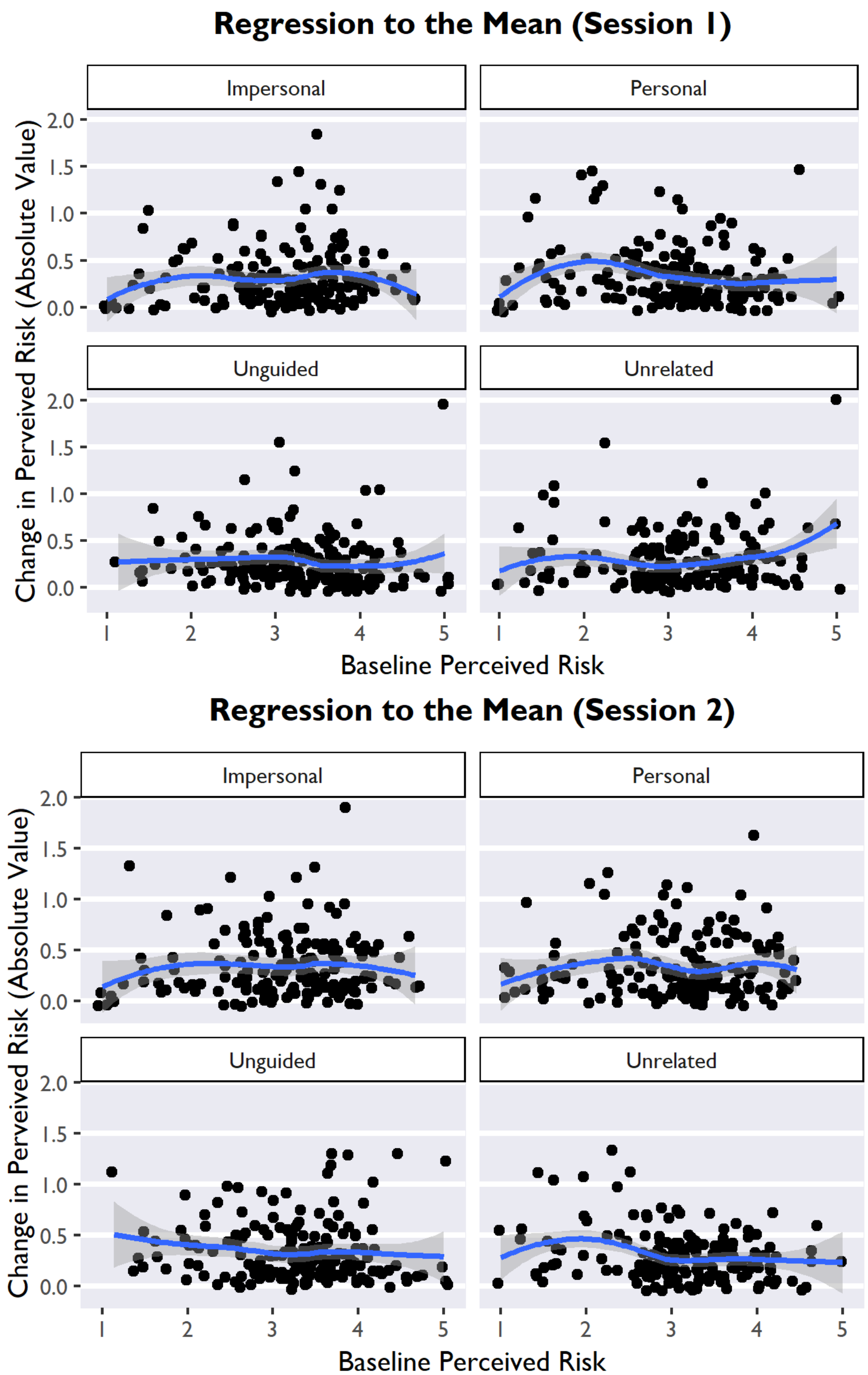

Figure S7. Scatterplot demonstrating the association between each participant's baseline perceived risk score and their post-intervention change in perceived risk, subset by session and intervention condition. If regression to the mean is present, then participants who reported extreme values of perceived risk at baseline (x-coordinates of 1 or 5) should demonstrate more change (greater y-values). This pattern would produce a U-shaped function. These plots do not reveal a U-shaped function, demonstrating that participants tended to report similar amounts of change in perceived risk, regardless of their baseline measures.

Results without Winsorizing

As described in the Methods section of the main text, we winsorized several variables in order to account for leptokurtic distributions. However, the significance of our results remained unchanged when variables were not winsorized. Below, we report alternate results with unaltered variables.

First, we tested for overall effects averaged across all intervention conditions. At the end of Session 1, there was a small average increase in perceived risk after the intervention, $t(734) = 3.85$, $p < .001$, Cohen's $d = 0.14$, 95% CI [0.07, 0.21]. This effect persisted to Session 2, $t(734) = 3.75$, $p < .001$, Cohen's $d = 0.13$, 95% CI [0.05, 0.21]. At the end of Session 1, there was also a moderate average decrease in willingness to engage in risky behaviors after the intervention, $t(734) = -13.37$, $p < .001$, Cohen's d = -0.49, 95% CI [-0.57, -0.42]. This effect persisted in Session 2, $t(672) = -5.39$, $p < .001$, Cohen's d = -0.21, 95% CI [-0.28, -0.13].

Next, we compared the effect of prediction error across the three simulation conditions. Using multiple linear regression, we found that average signed prediction error was weakly-to-moderately associated with change in perceived risk, $\beta = 0.20$, $t = 4.91$, $p < .001$, 95% CI [0.12, 0.28]. There was also an interaction between prediction error and simulation condition predicting change in perceived risk (Impersonal vs. Unrelated: $\beta = 0.14$, $t = 2.36$, $p = .019$, 95% CI [0.02, 0.25]; Impersonal vs. Personal: $\beta = -0.01$, $t = 0.58$, $p = .845$, 95% CI [-0.12, 0.10]). This interaction indicated that prediction error was significantly associated with change in perceived risk in the Impersonal Simulation condition ($r(175) = 0.35$, $p < .001$, 95% CI [0.21, 0.47]) and Personal Simulation condition ($r(176) = 0.18$, $p = .015$, 95% CI [0.04, 0.32]), but not in the Unrelated Simulation condition ($r(180) = 0.09$, $p = .246$, 95% CI [-0.06, 0.23]). Note that the variable for change in perceived risk was not winsorized for Session 2, because the distribution was approximately normal.

Next, we conducted the same analysis for an additional dependent variable: Change in willingness to engage in risky behaviors. Prediction error experienced during the Risk Estimation task was weakly negatively related to change in willingness, $\beta = -0.10$, $t = 2.51$, $p = .012$, 95% CI [-0.18, -0.02]. However, the interaction between prediction error and simulation condition was not significantly related to change in willingness (Impersonal vs. Unrelated: $\beta = -0.05$, $t = -0.87$, $p = .383$, 95% CI [-0.16, 0.06]; Impersonal vs. Personal: $\beta = -0.04$, $t = 0.63$, $p = .528$, 95% CI [-0.15, 0.08]). In Session 2, prediction error was not significantly related to lasting change in willingness to engage in risky activities, $\beta = -0.03$, $t = -0.62$, $p = .538$, 95% CI [-0.12, 0.06].

Controlling for Individual Differences

We tested whether the effects of prediction error and simulation condition held after controlling for several possible covariates: political conservatism, age, subjective numeracy ability (SNS), episodic future thinking ability (SAM-Future), self-reported vividness of the episodic simulation, and self-reported change in affect after the episodic simulation. In Session 1, the effect of prediction error on change in perceived risk ($\beta = 0.20$, $t = 4.57$, $p < .001$, 95% CI

[0.12, 0.29]) and the interaction between prediction error and simulation condition (Impersonal vs. Unrelated: $\beta = 0.15$, $t = 2.48$, $p = .014$, 95% CI [0.03, 0.27]) remained significant after controlling for demographic and individual difference variables. Conservatism was significantly positively associated with change in perceived risk ($\beta = 0.09$, $t = 2.42$, $p = .016$, 95% CI [0.02, 0.17]), likely because conservatives were more likely to underestimate risk at baseline ($r(733) = -0.36$, $p < .001$, 95% CI [-0.43, -0.30]). No other covariates showed a significant association with change in perceived risk.

As stated in our preregistration, we also tested whether any of the following variables interacted with prediction error and/or condition to predict change in perceived risk: vividness during the episodic simulation, change in affect after the simulation, subjective surprise ratings, subjective numeracy ability (SNS), and episodic future thinking ability (SAM-Future). We tested each of these variables in separate models and did not find any significant interaction effects. In the ANOVA tables below, we show that these covariates were not involved in any significant interactions, and our primary results (main effect of prediction error and interaction with condition) remained unchanged.

Table S2: Vividness ratings (from episodic simulation)

Row	*Sum.Sq*	*Df*	*F.value*	*P*
(Intercept)	0.05	1	0.05	0.82
Prediction Error	26.02	1	26.86	0.00***
Condition	0.26	2	0.13	0.88
Vividness	3.14	1	3.25	0.07
Prediction Error * Condition	8.38	2	4.33	0.01**
Prediction Error * Vividness	0.04	1	0.04	0.83
Condition * Vividness	1.26	2	0.65	0.52
Prediction Error * Condition * Vividness	0.42	2	0.22	0.80
Residuals	508.61	525		

Table S3: Affect ratings (from episodic simulation)

Row	*Sum.Sq*	*Df*	*F.value*	*P*
(Intercept)	0.06	1	0.06	0.80
Prediction Error	28.26	1	29.11	0.00***
Condition	0.31	2	0.16	0.85

Affect	0.93	1	0.96	0.33
Prediction Error * Condition	8.26	2	4.26	0.01**
Prediction Error * Affect	0.60	1	0.62	0.43
Condition * Affect	1.10	2	0.57	0.57
Prediction Error * Condition * Affect	0.33	2	0.17	0.85
Residuals	509.61	525		

Table S4: Subjective surprise ratings (from risk estimation task)

Row	*Sum.Sq*	*Df*	*F.value*	*P*
(Intercept)	0.16	1	0.17	0.68
Prediction Error	26.18	1	26.96	0.00***
Condition	0.14	2	0.07	0.93
Surprise	2.03	1	2.09	0.15
Prediction Error * Condition	5.44	2	2.80	0.06~
Prediction Error * Surprise	0.01	1	0.01	0.93
Condition * Surprise	0.44	2	0.23	0.80
Prediction Error * Condition * Surprise	0.18	2	0.09	0.91
Residuals	509.69	525		

Table S5: Subjective Numeracy Scale

Row	*Sum.Sq*	*Df*	*F.value*	*P*
(Intercept)	0.19	1	0.20	0.66
Prediction Error	23.80	1	24.63	0.00***
Condition	0.35	2	0.18	0.83
SNS	2.70	1	2.79	0.10
Prediction Error * Condition	7.65	2	3.96	0.02*
Prediction Error * SNS	0.00	1	0.00	0.97
Condition * SNS	1.52	2	0.79	0.46

Prediction Error * Condition * SNS	0.89	2	0.46	0.63
Residuals	507.39	525		

Table S6: Survey of Autobiographical Memory – Future Subscale

Row	*Sum.Sq*	*Df*	*F.value*	*P*
(Intercept)	0.18	1	0.18	0.67
Prediction Error	28.42	1	29.30	0.00***
Condition	0.13	2	0.06	0.94
SAM-Future	2.40	1	2.48	0.12
Prediction Error * Condition	8.09	2	4.17	0.02*
Prediction Error * SAM-Future	0.33	1	0.34	0.56
Condition * SAM-Future	0.25	2	0.13	0.88
Prediction Error * Condition * SAM-Future	0.26	2	0.13	0.88
Residuals	509.27	525		

Retrospective Report of Risky Activities

During Session 2 of Study 2, we asked participants to retrospectively report whether they had actually engaged in any of the 15 activities on the perceived risk scale during the delay period. The average number of activities reported was 1.67 (SD = 1.38). We conducted linear regression to predict the number of activities from the intervention condition (Personal, Impersonal, Unrelated, or Unguided) and the duration of the delay period. We found that there were no significant differences among intervention condition ($F_{(3,666)} = 0.81$, $p = .487$), but the delay length was positively associated with the number of activities reported ($\beta = 0.08$, $t = 2.06$, $p = .040$, 95% CI [0.004, 0.16]). Overall, our delay period was relatively short (M = 7.53 days, SD = 2.17) and participants likely did not have the opportunity to actually engage in many of the activities on our list.

Alternative Measure of Perceived Risk

One of the fifteen items on our perceived risk scale was “flying on an airplane.” Air travel may involve close contact with people from one’s local community (e.g., fellow passengers), but could also include people from surrounding counties and other cities (e.g., in the airport). To ensure that this ambiguity did not influence our results, we also calculated an alternative measure of perceived risk (average of 14 items instead of 15) that omitted the item for "flying on an airplane." We also applied the same change to the scale for willingness to engage in risky activities. Here, we redo the same analyses with the alternative scale, and show that this change to the scale makes no appreciable difference to our key findings.

Session 1: There was a significant main effect of prediction error on change in perceived risk, $\beta = 0.22$, $t = 5.12$, $p < .001$, 95% CI [0.13, 0.30]. There was also a significant interaction between prediction error and simulation condition predicting change in perceived risk (Impersonal vs. Unrelated: $\beta = 0.17$, $t = 2.77$, $p = .006$, 95% CI [0.05, 0.29]; Unrelated vs. Personal: $\beta = -0.17$, $t = -2.78$, $p = .006$, 95% CI [-0.29, -0.05]; Personal vs. Impersonal: $\beta = 0.01$, $t = 0.03$, $p = .973$, 95% CI [-0.12, 0.12]).

There was a significant main effect of prediction error on change in willingness to engage in risky activities, $\beta = -0.15$, $t = -3.51$, $p < .001$, 95% CI [-0.24, -0.07]. The interaction between prediction error and simulation condition was not significantly related to change in willingness (Unrelated vs. Impersonal: $\beta = -0.02$, $t = -0.33$, $p = .744$, 95% CI [-0.14, 0.10]; Impersonal vs. Personal: $\beta = -0.02$, $t = -0.40$, $p = .590$, 95% CI [-0.15, 0.10]; Unrelated vs. Personal: $\beta = 0.05$, $t = 0.78$, $p = .436$, 95% CI [-0.07, 0.17]).

Session 2: There was a significant main effect of prediction error on change in perceived risk, $\beta = 0.18$, $t = 4.25$, $p < .001$, 95% CI [0.10, 0.27]. The interaction between prediction error and simulation condition was not significantly related to change in perceived risk (Unrelated vs. Impersonal: $\beta = -0.11$, $t = -1.77$, $p = .077$, 95% CI [-0.23, 0.01]; Personal vs. Unrelated: $\beta = 0.09$, $t = 1.51$, $p = .131$, 95% CI [-0.03, 0.21]; Personal vs. Impersonal: $\beta = 0.02$, $t = 0.29$, $p = .771$, 95% CI [-0.10, 0.14]).

The main effect of prediction error on change in willingness to engage in risky activities was not significant, $\beta = -0.07$, $t = -1.54$, $p = .124$, 95% CI [-0.16, 0.02]. The interaction between prediction error and simulation condition was not significantly related to change in willingness (Unrelated vs. Impersonal: $\beta = -0.04$, $t = -0.67$, $p = .503$, 95% CI [-0.17, 0.08]; Impersonal vs. Personal: $\beta = 0.01$, $t = 0.19$, $p = .849$, 95% CI [-0.11, 0.14]; Unrelated vs. Personal: $\beta = 0.03$, $t = 0.49$, $p = .624$, 95% CI [-0.09, 0.16]).

Episodic Simulation Text

Below, we have reproduced the text used to guide participants through the three episodic simulation conditions. The three simulation conditions were matched in length and format. Participants were instructed to imagine each step of the episode and then type out the details that they visualized to confirm participation in the task. To encourage vivid imagining and thorough responses, participants were not allowed to advance to the next stage of the simulation until a minimum of 10 seconds had passed at each stage. Each of the three simulation conditions took approximately 5 minutes to complete. After the simulation task, participants rated their change in affect ("Overall, how do you feel after imagining this scenario?" 1 = *Much worse* … 5 = *Much better*) and subjective vividness of the simulation ("Overall, how vivid (clear and detailed) was the scene that you imagined?" 1 = *Not vivid at all* … 5 = *Extremely vivid*).

Personal Simulation. "In the next part of the study, you will imagine an event that could happen in your own life. We will guide you through the imagination exercise on the following pages. First, please think about four people who you know personally who you might invite over to your home for dinner. Please type in the box below to indicate the five people you chose (e.g., "my friend Martin", "my sister", "my boss"). You may use first names if you want, but to protect privacy, please do not write full names." [text entry]

"Now, please try to imagine what each of the four people look like when they are in your home. Close your eyes and try to visualize what their faces look like, what clothes they are wearing, and how it makes you feel to see them. When you are done imagining, please briefly describe what each person looks like or what they are wearing (1 sentence per person)." [text entry]

"Next, imagine the part of your house where you would be serving dinner. Close your eyes and try to visualize what the scene looks like, where each of your guests would be sitting, and what you would serve for dinner. When you are done imagining, please briefly describe what the room looks like and what you are eating for dinner. (2-3 sentences)" [text entry]

"Next, choose one of your four guests (other than yourself). Type below to indicate which guest you chose:" [text entry] "Imagine that the guest you chose begins coughing during dinner. They say that it may just be allergies. Close your eyes again and imagine what this scene would look like and how your other guests might react. Imagine how you would feel. Please describe how you would feel or what you would do in the box below. (1-2 sentences)" [text entry]

"Now, imagine that three days later, the guest tells you that they have tested positive for COVID-19 and are going to the hospital because they feel very sick. Imagine that you have to contact each of your other guests and tell them that they may have been infected at your home. Think about what it would be like to talk to each of your guests, what you would say to them, and the emotions that you would feel. Please describe how you would feel in the box below. (1-2 sentences)" [text entry]

"Imagine that you also become sick with COVID-19 after your dinner party. You have a fever, feel dizzy, and have a cough that makes it hard to breathe. Close your eyes and try to imagine what these symptoms would feel like. Please describe how you would feel in the box below. (1-2 sentences)" [text entry]

Impersonal Simulation. "In the next part of the study, you will imagine an event that could happen in someone's life. We will guide you through the imagination exercise on the following pages. First, please think about a man named Martin and four people who he has invited to a dinner party at his house. Martin's other guests are his wife, coworker, and two of their friends. Please make up names for Martin's four guests, and type their names in the boxes below." [text entry]

"Now, please try to imagine what Martin and each of the other four people look like when they are at the dinner party. Close your eyes and try to visualize what their faces look like, what clothes they are wearing, and what they are feeling. When you are done imagining, please briefly describe what each person looks like or what they are wearing (1 sentence per person)." [text entry]

"Next, imagine the part of Martin's house where the party guests are seated for dinner. Close your eyes and try to visualize what the scene looks like, where each of the people would be sitting, and what they are eating for dinner. When you are done imagining, please briefly describe what the room looks like and what the people are eating for dinner. (2-3 sentences)" [text entry]

"Next, choose one of the four party guests (other than Martin). Type below to indicate which person you chose:" [text entry] "Imagine that the guest you chose begins coughing during dinner. They say that it may just be allergies. Close your eyes again and imagine what this scene would look like and how the other guests might react. Imagine how they would feel. Please describe how the party guests would feel or what they would do in the box below. (1-2 sentences)" [text entry]

"Now, imagine that three days later, the guest tells Martin that they have tested positive for COVID-19 and are going to the hospital because they feel very sick. Imagine that Martin has to contact each of the other guests and tell them that they may have been infected at his home. Think about what it would be like for Martin to talk to each of the guests, what he would say to them, and the emotions that he would feel. Please describe how Martin would feel in the box below. (1-2 sentences)" [text entry]

"Imagine that Martin also becomes sick with COVID-19 after the dinner party. Martin has a fever, feels dizzy, and has a cough that makes it hard to breathe. Close your eyes and try to imagine Martin experiencing these symptoms. Please describe how Martin would feel in the box below. (1-2 sentences)" [text entry]

Unrelated Simulation. "In the next part of the study, you will imagine an event. We will guide you through the imagination exercise on the following pages. First, please imagine a rabbit named Martin, who lives together with four other rabbits. Please make up names for the four other rabbits who live with Martin, and type their names in the box below." [text entry]

“Now, please try to imagine what Martin and the other four rabbits look like. Close your eyes and try to visualize what their bodies look like, what color fur they have, and what they are doing. When you are done imagining, please briefly describe what each rabbit looks like (1 sentence per rabbit).” [text entry]

“Next, imagine that Martin finds a vegetable garden in a backyard, and brings his friends there to find food. Close your eyes and try to visualize what the scene looks like, where each of the rabbits is sitting, and what vegetables are in the garden. When you are done imagining, please briefly describe what the garden looks like and what vegetables the rabbits are eating. (2-3 sentences)” [text entry]

“Next, choose one of the rabbits (other than Martin). Type below to indicate which rabbit you chose.” [text entry] “Imagine that the rabbit you chose discovers that the vegetables are rotten, and warns the other rabbits that the vegetables they have been eating might not be safe. Imagine what this scene would look like and how the other rabbits might react. Imagine how they would feel. Please describe how the rabbits would feel or what they would do in the box below. (1-2 sentences)” [text entry]

“Now, imagine that three days later, the rabbit you chose is feeling very sick after eating the rotten vegetables. Imagine that Martin has to tell the other rabbits that their friend is sick because the garden he found was full of bad vegetables. Think about what it would be like for the rabbits and what they would feel. Please describe how Martin would feel in the box below. (1-2 sentences)” [text entry]

“Imagine that Martin also becomes sick after eating the bad vegetables. Martin cannot eat and feels very weak. Close your eyes and try to imagine Martin experiencing these symptoms. Please describe how Martin would feel in the box below. (1-2 sentences)” [text entry]

References

1. D. Lüdecke, *sjPlot: Data Visualization for Statistics in Social Science* (2021).

2. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag New York, 2016).

3. D. Lakens, Equivalence Tests: A Practical Primer for t Tests, Correlations, and Meta-Analyses. *Soc. Psychol. Personal. Sci.* **8**, 355–362 (2017).

4. COVID-19 Activity Risk Index. *COVID-19 RECoVERY Consult.* (May 25, 2021).