Original Manuscript

Social Psychological and Personality Science
1–12

DOI: 10.1177/19485506211056761
journals.sagepub.com/home/spp
SAGE

Risky Therefore Not Beneficial: Replication and Extension of Finucane et al.'s (2000) Affect Heuristic Experiment

Emir Efendić[1]*, Subramanya Prasad Chandrashekar[2]*, Cheong Shing Lee[3]*, Lok Yan Yeung[3]*, Min Ji Kim[3]*, Ching Yee Lee[3]*, and Gilad Feldman[3]

Abstract

Risks and benefits are negatively related in people's minds. Finucane et al. causally demonstrated that increasing risks of a hazard leads people to judge its benefits as lower. Vice versa, increasing benefits leads people to judge its risks as lower (original: $r = -.74$ [−0.92, −0.30]). This finding is consistent with an affective explanation, and the negative relationship is often presented as evidence for an affect heuristic. In two well-powered studies, using a more stringent analytic strategy, we replicated the original finding. We observed a strong negative relationship between judgments of risks and benefits across three technologies, although we do find that there was no change in risks when highlighting low benefits. We note that risks seem to be more responsive to manipulation (as opposed to benefits) and find evidence that the negative relationship can depend on incidental mood. We provided materials, data sets, and analyses on https://osf.io/sufjn/?view_only=6f8f5dc6ff524149a4ed5c6de9296ae8.

Keywords

affect heuristic, judgment and decision-making, heuristics, risk, replication

Introduction

People tend to view risks and benefits as negatively related: the riskier something is, the less beneficial it is. However, risks and benefits are distinct concepts and are sometimes even positively correlated—some technologies or hazards that are beneficial may be high or low in risk, but those that are not beneficial are unlikely to be high in risk. In a seminal article, Finucane et al. (2000) proposed that the negative relationship occurs due to an *affect heuristic* (AH) whereby people rely on affect when judging the risks/benefits of specific hazards. Furthermore, they demonstrated evidence that is consistent with an affective explanation of this relationship. Take nuclear energy for example. The AH proposes that increasing the risks of nuclear energy (e.g., by exalting the hazard uranium has for human health) turns the affective evaluation associated with it negative, thereby leading people to judge its benefits as lower. Vice versa, increasing benefits leads to positive affect and to people judging its risks as lower (see Table 1).

Affect Heuristic

Affect is a crucial component of people's decision-making (Kahneman, 2003, 2011; Lerner et al., 2015; Loewenstein et al., 2001; Rachlin, 2003). It is argued that reliance on affect is often a much quicker, easier, and more efficient way to navigate the complexities of everyday decision-making (Damasio, 1994; Schwarz & Clore, 1983) and that affect informs many judgments and decisions (Albarracín & Kumkale, 2003; Peters et al., 2006; Schwarz, 2012; Slovic et al., 2002; Wyer et al., 1999).

Early studies of risk perception have shown that feelings of dread are major determinants of public perception and acceptance of risk for a wide range of hazards (Slovic, 1987). Focusing on this link, Finucane et al. (2000) proposed that people use an *affect heuristic* (AH) when making risk judgments. According to this view, people may use their affective response to a risk (e.g., "How do I feel about nuclear energy?") to infer how large they consider the risk to be. The argument is that: "Using an overall, readily

[1]Maastricht University, School of Business and Economics, Department of Marketing and Supply Chain Management, the Netherlands
[2]Hong Kong Metropolitan University, Hong Kong
[3]The University of Hong Kong, Hong Kong
*Contributed equally, joint first authors.

Corresponding Author:
Gilad Feldman, Department of Psychology, The University of Hong Kong, Pok Fo Lam Road, Hong Kong 999077.
Email: gfeldman@hku.hk

Table 1. Summary of the Predictions According to the Affect Heuristic (AH).

Manipulated attribute	Impact on affect	Impact on non-manipulated attribute
Risk is high	Negative affect	Benefit is low
Risk is low	Positive affect	Benefit is high
Benefit is high	Positive affect	Risk is low
Benefit is low	Negative affect	Risk is high

available affective impression can be far easier—more efficient—than weighing the pros and cons or retrieving from memory many relevant examples, especially when the required judgment or decision is complex or mental resources are limited" (Finucane et al., 2000, p. 3).

Reliance on affect is a general process and, consistent with an AH, a wide range of findings support the idea that affect provides valuable information that people use to simplify their decision-making. For instance, affect-laden imagery has been shown to predict people's preferences in investment decisions (MacGregor et al., 2000), smoking (Benthin et al., 1995), information integration (Anderson, 1981; Efendić et al., 2019), simple choice gambles (Bateman et al., 2007), and morality judgments (Slovic & Västfjäll, 2010).

Risks and Benefits

For a long time, the negative relationship between judgments of risks and benefits puzzled researchers (Fischhoff et al., 1978) as these judgments should be positively correlated or independent of one another (Slovic, 1987). In a breakthrough study, Alhakami and Slovic (1994) found that the negative relationship was linked to how a person generally feels about a hazard. Later, Finucane et al. (2000) showed that the inverse relationship between risk and benefits was strengthened under time pressure designed to limit analytic thinking (their Study 1) and that it is causally determined. Specifically, manipulating one attribute—for example, increasing risk—led to an affectively congruent but inverse relationship, that is, decreased benefit and vice versa (their Study 2).

This inverse relationship has been observed elsewhere as well. It has been found that when general negative affect is evoked (i.e., participants were shown photographs depicting houses in flooded regions), this led to increased levels of perceived risk (Keller et al., 2006). Similarly, incidental negative affect (e.g., negative mood) was found to amplify reliance on affect, which led to stronger negative correlations between risks and benefits (Västfjäll et al., 2014). Interestingly, affective association with a particular hazard has been shown to influence the interpretation of new information. People evaluated nuclear power more negatively than solar power because of more negative feelings associated with nuclear power (Siegrist & Sütterlin, 2014). Similar negative associations between risk and benefits have been found in consumer judgments of novel products (King & Slovic, 2014), in the financial domain (Ganzach, 2000), and in wood smoke pollution (Bhullar et al., 2014). Recently, Skagerlund et al. (2020) found that the negative correlation is tied to cognitive reflection ability.

Replication Value and Present Research

In this article, with two well-powered studies, we aimed to closely replicate and extend our understanding of the causal demonstration of the negative relationship between risks and benefits, using the same materials and procedure as in the original paper (Finucane et al., 2000).

We chose to replicate Study 2 from Finucane et al. (2000) for several reasons. First, while many correlational studies have found the negative relationship, few demonstrated it causally. King and Slovic (2014) used a similar method as Finucane and colleagues, but other work mostly found correlational support (some research has even failed to find the same relationship, Raue et al., 2019). There is therefore value in demonstrating, with sufficient statistical power, whether the causal effect is robust. Second, the analysis approach used in the original studies and in later demonstrations of the negative relationship (e.g., King & Slovic, 2014) were nonstandard, failing to account for non-independence of data and relying on counting the number of times the manipulation worked in the predicted direction—a strategy that leads to large information loss. A more stringent analytic approach with mixed-effect modeling ought to provide information on the generalizability of the effect. Third, the findings are relevant for risk communication. Changing risk/benefit judgments by manipulating solely one attribute (either risk or benefit) has vast applied potential. Risk campaigns can focus on changing people's judgments about many plights of today's society (e.g., smoking, obesity, and so on). Fourth, as of this writing, we are unaware of any other attempts to directly replicate this study. This is surprising given the relevance in understanding the relationship between risks and benefits, as well as the popularity of the original article and how it promoted the AH in the judgment and decision-making literature. As of this writing, the original article has been cited 3,363 times with a later updated review article being cited 3,860 times (Slovic et al., 2007).

We also wish to highlight an important distinction. The observation of the negative relationship is often presented as evidence for an AH in risk judgments. For example, observing the negative relationship leads authors to conclude that the AH is a robust phenomenon (Skagerlund et al., 2020). However, the original, as well as many other studies, fail to demonstrate that it is affect that mediates this relationship (although converging evidence on the importance of affect would suggest this is the case). Our aim here is to replicate the negative causal relationship between risks and benefits. As such, this replication also does not speak to the mechanism that underlies the relationship. Other more cognitive, rather than affective, mechanisms remain a plausible explanation. Nevertheless, we hope that investigating whether the causal relationship replicates will (a) provide important insight into this interesting phenomenon and (b) serve other researchers who wish to use the paradigm to further understand whether it is affect or something else that explains it.

We thus consider this investigation to be a needed direct replication. Replications should be sufficiently similar to the original study to adequately gauge support for the original findings (LeBel et al., 2019). Furthermore, given the prevalence of publication bias (Bakker et al., 2012), a close replication adds value by providing evidence that strengthens or weakens the finding.

Overview of Studies

This replication was part of an ongoing replications project (see Supplementary Figure S1 and the project process section in the supplementary material for more details). We crowdsourced the replication using two teams, both teams being supervised by experienced authors. Each team collected data independently and wrote detailed preregistrations. We thus report the results of two studies serving as close replications of Study 2 from Finucane et al. (2000), using the same methodology and the same materials.[1] The two studies differ only in the target sample, one obtained on MTurk (U.S. participants) and the other on Prolific (U.K. participants). The two studies were preregistered on the OSF (MTurk: https://osf.io/ab5dw/files/?view_only=9084199ffd9a4dfdbd0e80765a206bc6; Prolific: https://osf.io/p4qjx/files/?view_only=9fbbe4f1ebcd444988f74b1a6ad31346).[2] All materials, data sets, and analysis scripts are available on OSF (https://osf.io/sufjn/?view_only=6f8f5dc6ff524149a4ed5c6de9296ae8). We report how we determined the sample size, all data exclusions (if any), all manipulations, and all measures.

Extensions

In addition to the direct replication of Study 2 from Finucane et al. (2000), we also report two extensions. First, we looked at the effect of naturally occurring incidental mood on the negative relationship between judgments of risks and benefits. In the MTurk sample, participants were asked to rate their current levels of (a) pleasure—*unpleasant* vs. *pleasant* and (b) arousal—*deactivated* vs. *activated* (using two affective sliders that ranged from −100 to 100, centered in the middle). We based our measure on core affect that represents states experienced as simply feeling good or bad, energized, or enervated (Russell, 2003). We use the term "naturally occurring incidental mood" to highlight that this is a measured rather than manipulated variable and that the affect in question is incidental (i.e., unrelated to the judgment at hand). Any affect that arises due to changes in risk/benefit descriptions is integral (i.e., affect stemming from the judgment target at hand). Several predictions can be made on how naturally occurring incidental mood could impact the negative relationship: (a) incidental mood is misattributed (Schwarz, 2012) to risk/benefit judgments impacting the strength of the negative correlations, (b) incidental affect has a specific effect in that negative incidental affect leads to high risk and low benefit, while positive incidental affect leads to low risk and high benefit, not impacting the strength of the negative correlations; or (c) it has a negation effect where, akin to mood regulation models for example (Andrade, 2005), being in a pleasurable naturally occurring mood may interfere with people's ability to effectively map a negative change in integral affect (e.g., by describing risks as high). Highlighting the interaction between such incidental and integral states can offer insights into the role of affect in the negative relationship.

Second, we explored whether there was a stronger negative relationship when risks, as opposed to benefits, are manipulated. Illuminating this boundary condition could provide insight into which of these two attributes people find more informative or important for their risk judgments.

Method

Participants

In the first study, a total of 806 participants from the United States were recruited through MTurk using the TurkPrime platform (Litman et al., 2017). In the second, a total of 1,008 participants from the United Kingdom were recruited through Prolific. To determine the number of participants needed, we conducted a power analysis planning to detect the weakest effect size reported in the original *that was also significant* (at $p < .05$). Therefore, given our resource constraints, we based our power analysis on having 95% power to detect a Cohen's $d_z = 0.30$. This resulted in a suggested sample size of 147 participants per condition and a total of 588 across 4 between-subject conditions. Finally, we aimed for a higher sample size between 750 and 800 participants, as this would also ensure we were able to detect a smaller effect size (Cohen's d_z) of .20 at 80% power. A comparison of the target article sample and the replication samples is provided in Table S1 in the supplementary material.

To obtain the final sample, we first excluded (30 from MTurk sample and 40 from Prolific sample) participants following our preregistered exclusion criteria.[3] Because the studies were identical, we combined[4] them for the final data analysis, resulting in 1,552 participants (*MTurk* = 776; *Prolific* = 776; M_{Age} = 38.99, SD_{Age} = 12.30; 822 females, 727 males, 3 would rather not say).

Design, Procedure, and Measures

Both studies had a 2 (Between-subject factor—Direction: High vs. Low) × 2 (Between-subject factor—Manipulated Attribute: Risk vs. Benefit) × 3 (Within-subject factor—Technology Scenario: Nuclear Power vs. Natural Gas vs. Food Preservative) mixed-subject design (see Table S3 and Table S4 in the supplementary material for more details and full descriptions of the measures and direction/attribute information). Please note that the second study (Prolific) included an additional experimental condition that was excluded due to a methodological issue.[5]

Participants were first asked to answer questions regarding the perceived benefit and risk of all three technologies (Nuclear Power, Natural Gas, Food Preservatives)—the same ones used in the original study. The presentation of the technologies was randomized. Participants were asked two questions, in random order, for each technology, namely: "In general, how risky [beneficial] do you consider the use of nuclear power / natural gas / food preservatives?"[6], answering on a 10-point scale from 1 (*not at all risky [beneficial]*) to 5 (*moderate risk [benefit]*) to 10 (*very risky [beneficial]*).

Subsequently, dependent on the conditions, participants were presented with textual vignettes designed to change the affective quality (e.g., high risk = negative, high benefit = positive, and so on) of the scenarios. We used the same descriptions from the original study (https://osf.io/y97tp/). For example, in the low benefit condition for the hazard natural gas, participants were presented with the following text (shortened):

> Natural gas is used as a source of energy in the US. Natural gas has the property of being a gas at room temperature, which allows it to be burned to produce heat. However, this same gaseous property limits the energy tasks that natural gas can be used for. Natural gas is not able to replace electricity for such tasks as lighting, or the numerous jobs that need electric motors, such as refrigeration or the operation of machinery.

After reading the information, participants again provided answers to the risk and benefit questions for each technology scenario. Please note that once participants were assigned to one of the between-subject conditions, they were in that condition for all three scenarios, as the scenario was a within-subject variable. This means that we had data from 4,656 trials. Finally, participants answered a funneling section and provided demographic information. At the end of the study, a short debriefing was given regarding the study's purpose and confidentiality. We characterize the current replication as a "very close replication" based on the framework for classification of the replications using the criteria by LeBel et al. (2018; see Table S45 in the supplementary material).

Results

Analysis Strategy

We report both the original (i.e., repeating the same analytic strategy as in Finucane et al., 2000) and an improved analytic approach. For the improved, we employed linear mixed-effects models (LMEM) using the lme4 package in *R* (Bates et al., 2015). Significance for fixed effects was assessed via Satterthwaite's degrees of freedom (Kuznetsova et al., 2017). Unless stated otherwise, the models adjusted for covariates at Level 1 (ratings of risks and benefits before the experimental treatment) and Level 2 (i.e., Technology type and participants' ID were treated as random effects). We added pre-scores on the manipulated/nonmanipulated attribute to reduce noise of our assessment and to check whether the preratings may moderate the effect of the manipulation. LMEMs reduce the chance of Type I errors, account for nonindependence of data points (e.g., within-subject observations), provide a greater flexibility with specification of the covariance structure, and allow us to make more generalizable claims across samples of participants and stimuli (hazards in our case; Judd et al., 2012).

Original Data Analytic Approach (Finucane et al., 2000)

Descriptive statistics of the measures across the two samples are noted in Table S39 and Table S40 of the supplementary material. Following the original approach, we conducted paired samples t tests (two-tailed). Specifically, for each technology, we compared the mean pre- and post-manipulation ratings of the manipulated and the non-manipulated attributes. Positive t-values indicate that there was an increase in rating after manipulation. Negative t-values indicate there was a decrease in rating after manipulation. The results are in line with the original finding (See Table S41–S44 in the supplementary material for the detailed results). That is, for the manipulated attribute ratings, providing information on high and low benefits or risks led to higher and lower post-manipulation ratings of benefits or risks. For the non-manipulated attribute, we see the inverse: providing information on high and low benefits or risks led to lower and higher post-manipulation ratings of risk and benefits.

Furthermore, we tested the correlation between risk and benefits using the t-values from the abovementioned analysis. We found strong support for a negative correlation: MTurk sample: $r(10) = -.87$, 95% confidence

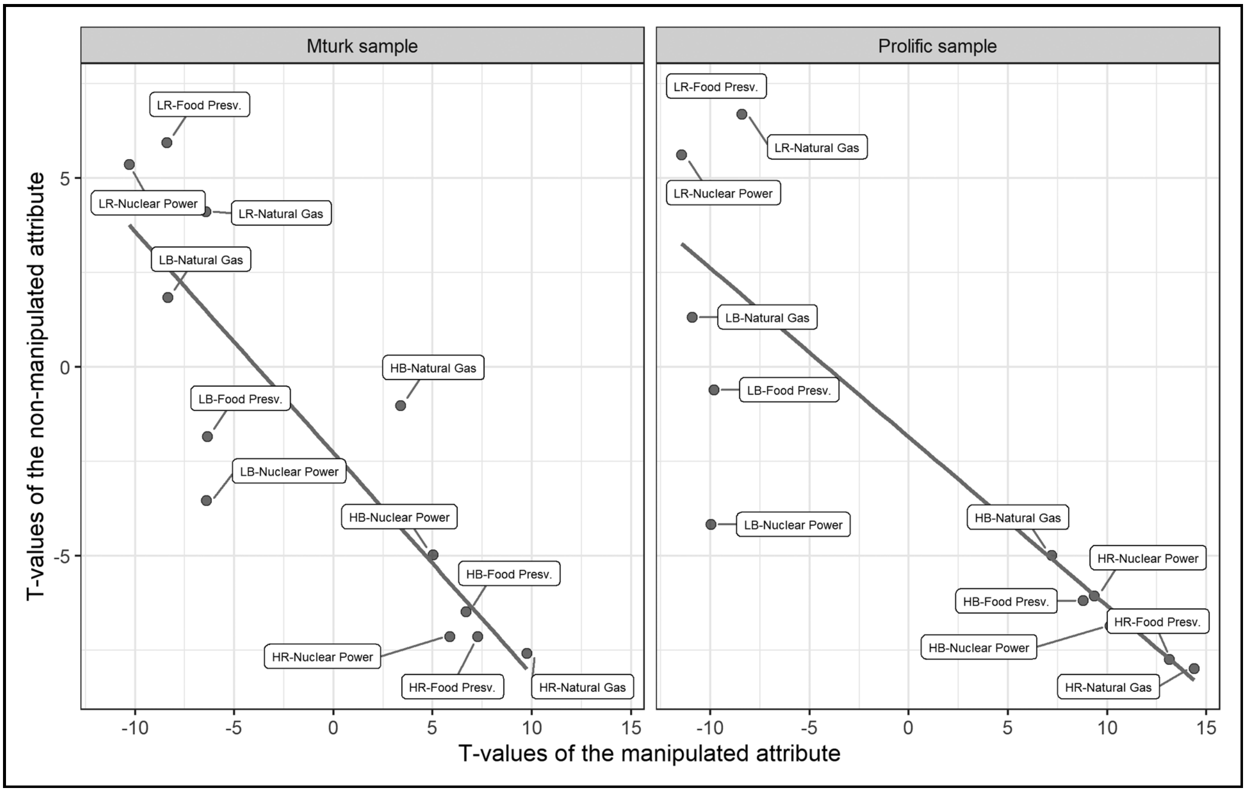

Figure 1. t-Values for Manipulated Versus Non-Manipulated Attributes.
Note. *t*-values for four-direction/attribute information manipulations (HB = High Benefit; LB = Low Benefit, HR = High Risk, LR = Low Risk) for the three technologies (nuclear power, natural gas, and food preservatives) across the two samples (MTurk and Prolific). The negative slope shows the predicted negative relationship between risks and benefits—as benefits increase risks decrease and as risks increase benefits decrease.

interval (CI): [−0.96, −0.59], $p = .003$; Prolific sample: $r(10) = -.84$, 95% CI = [−0.95, −0.50], $p < .001$. Plotting the *t*-values in Figure 1, the negative slope shows that when ratings on the manipulated attribute increase, ratings on the non-manipulated attribute decrease (and vice versa). Simply put, when benefits increase risks decrease and when risks increase benefits decrease, indicating a negative relationship.

Mixed-Model Approach

Manipulation Checks. We conducted LMEMs with change in the manipulated attribute as the DV (i.e., ratings on a manipulated attribute after experimental treatment minus ratings on manipulated attribute before experimental treatment; 0, therefore indicates no change, a positive value an increase, and negative value indicates a decrease). Table 2 presents the fixed-effects coefficients with all the predictors (See Table S11–S14 in the supplementary material for step-by-step regression results).

The significant effect of Direction shows that, regardless of the manipulated attribute, if the direction was high there was a positive change while if the direction was low there was a negative change, indicating a successful manipulation check (see Figure 2 and Tables S41–S44 for detailed statistics).

Negative Relationship Between Risks and Benefits. To test whether we observe a negative relationship between risks and benefits, we looked at the effects of the manipulated attribute on the nonmanipulated attribute. Specifically, we regressed change in ratings of nonmanipulated attributes (DV) on Direction, Manipulated Attribute, and their interaction, adjusting for covariates at Level 1 (Pre-rating manipulated attribute; and three-way interaction between pre-rating non-manipulated attribute, Direction, and Manipulated Attribute) and Level 2 (i.e., Technology type and participant's ID). Table 3 summarizes these results (see Table S20–S24 in the supplementary material for step-by-step regression results and model comparisons).

The main effect of direction supports the original finding of the negative relationship. In addition, we find that the directionality of pre- and post-treatment changes in the non-manipulated attribute was consistent with the predicted inverse relationship, except in the Low-benefit condition (see Figure 3 and Tables S41–S44 for detailed statistics).

Table 2. Estimated Fixed-Effects Coefficients of the Mixed-Effects Regression Model With Change in the Manipulated Attribute as the DV.

Predictors	DV: Change in non-manipulated attribute			
	B	*SE*	95% CI	*p*
Intercept	−0.09	0.06	[−0.21, 0.04]	.185
Pre-rating manipulated attribute (PMA)	−1.09	0.03	[−1.15, −1.03]	**<.001**
Direction (high vs. low)	2.56	0.07	[2.42, 2.69]	**<.001**
Manipulated attribute (risk vs. benefit)	−0.27	0.07	[−0.40, −0.13]	**<.001**
Direction × manipulated attribute	0.49	0.14	[0.22, 0.75]	**<.001**
PMA ×direction	−0.10	0.06	[−0.22, 0.02]	.109
PMA × manipulated attribute	0.01	0.06	[−0.11, 0.14]	.819
PMA × direction × manipulated attribute	0.16	0.12	[−0.08, 0.40]	.199

Note. Variables were coded as follows—direction: −0.5 = low, + 0.5 = high; attribute: −0.5 = benefit, + 0.5 = risk. CI = confidence interval.

Exploratory Analysis: Mediation Effects. We also tested whether the effect of the experimental manipulation on change in the non-manipulated attributes was mediated by the changes in the manipulated attribute as the analytic reasoning would suggest. To do this, we conducted a multilevel mediation analysis (this analysis was not part of the pre-registration). Bayesian estimation of the multilevel mediation model was performed using the *bmlm* R package (Vuorre & Bolger, 2018). Because our experimental design involved two directions (High vs. Low), we conducted two independent mediation analyses that looked at the responses within High and Low separately. Indeed, both sets of mediation analysis show a significant indirect effect of manipulation on non-manipulated attribute rating through manipulated attribute rating (High only mediation: $M_{posterior} = -0.54$, $SD = 0.04$, $CI = [-0.61, -0.47]$; Low only mediation: $M_{posterior} = 0.55$, $SD = 0.04$, $CI = [0.48, 0.62]$). For details results see Table S25–S26 in the supplementary material.

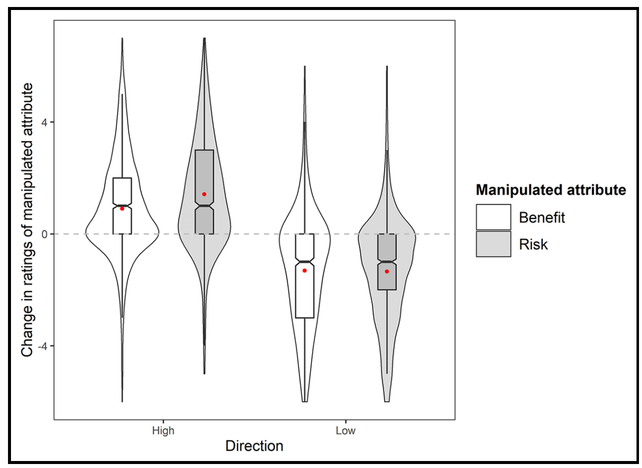

Figure 2. Distribution of Ratings on Change in Manipulated Attribute as DV by Experimental Conditions.
Note. Figure includes violin plots displaying the distribution of responses, boxplots displaying the median, first, and third quartiles, while the mean value is identified by the red circle.

Extensions

Naturally Occurring Incidental Mood and the Negative Relationship Between Risks and Benefits. We conducted an analysis where the change in ratings of manipulated attributes, level of pleasure, level of arousal, and their interaction were set as predictors of change in the ratings of the non-manipulated attributes. Table 4 and Figure 4 summarize the results. As a representation of the negative relationship between risks and benefits, we looked at predicting change in non-manipulated attribute with change in manipulated attribute. Indeed, a negative correlation between these two variables represents the negative relationship. We decided to use this (rather than an interaction between the dummy coded direction and manipulated attribute), as it is easier to represent and interpret a potential two-way interaction with pleasure or arousal.

We found some support that the negative relationship is moderated by incidental pleasure (see Figure 4). Specifically, the negative relationship was stronger among participants who reported higher incidental pleasure in comparison to participants who reported lower incidental pleasure.

Risk/Benefit Strength. We also examined whether there was a stronger negative relationship when risks, as opposed to benefits were manipulated and the extent to which it may depend on the manipulated conditions. For the analysis, similar to above, we again used the change in ratings of manipulated attributes, Manipulated Attribute (Risk vs. Benefit), Direction, and their interaction as predictors of change in the ratings of the non-manipulated attributes. Table 5 and Figure 5 summarize the results.

The interaction between manipulated attribute and CMA (change in manipulated attribute) indicates that the

Table 3. Estimated Fixed-Effects Coefficients of the Mixed-Effects Regression Model With Change in the Non-Manipulated Attribute as the DV.

Predictors	DV: Change in non-manipulated attribute			
	B	*SE*	95% CI	*p*
Intercept	−0.26	0.10	[−0.45, −0.06]	**.009**
Pre-rating manipulated attribute (PMA)	−0.21	0.03	[−0.27, −0.15]	**<.001**
Pre-rating non-manipulated attribute (PNMA)	−0.95	0.03	[−1.01, −0.89]	**<.001**
Direction (high vs. low)	−1.15	0.06	[−1.27, −1.03]	**<.001**
Attribute (risk vs. benefit)	0.55	0.06	[0.43, 0.67]	**<.001**
PNMA × Direction	0.14	0.05	[0.04, 0.25]	**.008**
PNMA × Attribute	−0.16	0.06	[−0.27, −0.05]	**.004**
Direction × Attribute	−1.34	0.12	[−1.58, −1.10]	**<.001**
PNMA × Direction× Attribute	0.13	0.11	[−0.08, 0.35]	.221

Note. Variables were coded as follows—direction: −0.5 = low, + 0.5 = high; attribute: −0.5 = benefit, + 0.5 = risk. CI = confidence interval.

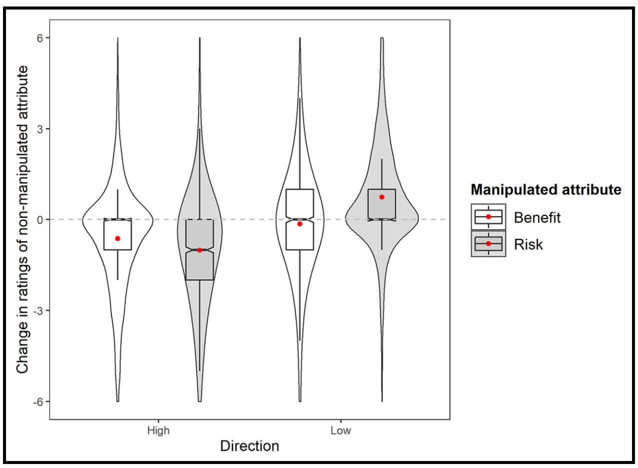

Figure 3. Distribution of Rating on Change in Non-Manipulated Attribute as DV by Experimental Conditions.
Note. Figure includes violin plots displaying the distribution of responses, boxplots displaying the median, first, and third quartiles, while the mean value is identified by the red circle.

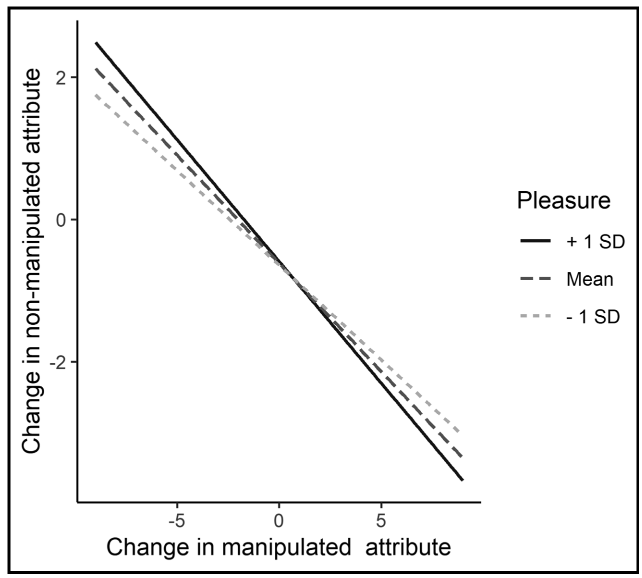

Figure 4. The Interaction Between Change in Manipulated Attribute and Pleasure on Change in Non-Manipulated Attribute

strength of the negative relationship between the manipulated and non-manipulated attribute was stronger when risks, as opposed to benefits, were manipulated. Furthermore, the three-way interaction (Direction × Manipulated Attribute × CMA) suggests that the extent of difference between risks and benefits varies as a function of the direction of manipulation (High vs. Low). Proceeding to conduct separate analyses for Low and High conditions, results within the high condition show no support for interaction. However, results within the low condition do find support for the interaction (See Table S32 and Table S33 in the supplementary material for detailed results). This lack of consistency leads us to conclude that the strength of the negative relationship between the manipulated and the non-manipulated attribute being stronger when risks, as opposed to benefits, were manipulated is mainly driven by participants' responses within the Low-Benefit condition (see Figure 5). Specifically, we note large differences in change in ratings of non-manipulated attribute across Risk, M_{change} = 0.74 (SE = 0.05) and Benefit, M_{change} = −0.14, (0.05), manipulation within the low condition. However, those differences are much smaller within the high condition, Risk: M_{change} = −1.01 (0.06); Benefit: M_{change} = −0.62, (0.05).

General Discussion

In two studies, using samples from the United States and the United Kingdom, we re-did Study 2 from Finucane et al. (2000). With high power and using a more precise analytic approach, we successfully replicated and obtained a similar effect as in the original study providing support

Table 4. Estimated Fixed-Effects Coefficients From the Mixed-Effects Regression Model Adding Pleasure and Arousal Measures on Change in Non-Manipulated Attribute as DV.

Predictors	DV: Change in non-manipulated attribute			
	B	*SE*	95% CI	*p*
Intercept	−0.59	0.15	[−0.88, −0.29]	<**.001**
Pre-rating non-manipulated attribute (PNMA)	−0.63	0.05	[−0.72, −0.54]	<**.001**
Pre-rating manipulated attribute (PMA)	−1.05	0.04	[−1.13, −0.97]	<**.001**
Pleasure	0.03	0.05	[−0.07, 0.13]	.557
Arousal	−0.06	0.05	[−0.16, 0.04]	.266
Change in manipulated attribute (CMA)	−0.70	0.05	[−0.79, −0.61]	<**.001**
Direction (high vs. low)	0.30	0.09	[0.13, 0.48]	**.001**
Manipulated attribute (risk vs. benefit)	0.39	0.08	[0.23, 0.56]	<**.001**
Pleasure × Arousal	−0.02	0.03	[−0.08, 0.04]	.536
Pleasure × CMA	−0.09	0.04	[−0.16, −0.01]	**.025**
Arousal × CMA	0.05	0.04	[−0.04, 0.13]	.293
Pleasure × Arousal × CMA	−0.03	0.03	[−0.09, 0.02]	.201

Note. CI = confidence interval.

Table 5. Estimated Fixed-Effects Coefficients From the Mixed-Effects Regression Model Looking at Moderation of the Negative Relationship by Risks/Benefits.

Predictors	DV: Change in non-manipulated attribute			
	B	*SE*	95% CI	*p*
Intercept	−0.24	0.13	[−0.49, 0.02]	.066
Pre-rating manipulated attribute (PMA)	−0.59	0.03	[−0.65, −0.53]	<**.001**
Pre-rating non-manipulated attribute (PNMA)	−1.01	0.03	[−1.06, −0.95]	<**.001**
Direction (high vs. low)	−0.37	0.06	[−0.49, −0.24]	<**.001**
Manipulated attribute (risk vs. benefit)	0.44	0.06	[0.31, 0.56]	<**.001**
Change in manipulated attribute (CMA)	−0.74	0.03	[−0.80, −0.68]	<**.001**
Direction × Manipulated Attribute	−0.85	0.12	[−1.09, −0.60]	<**.001**
Direction × CMA	−0.12	0.05	[−0.23, −0.02]	**.022**
CMA × Manipulated Attribute	−0.27	0.05	[−0.37, −0.16]	<**.001**
Direction × Manipulated Attribute × CMA	0.22	0.11	[0.01, 0.43]	**.037**

Note. Variables were coded as follows—direction: −0.5 = low, + 0.5 = high; attribute: −0.5 = benefit, + 0.5 = risk. CI = confidence interval.

for the demonstration of a causal negative relationship between risks and benefit judgments. Specifically, we showed that increasing the risks of three technologies (nuclear energy, food preservatives, and natural gas) led to lower judgments on benefits while increasing the benefits led to lower judgments on risks. Vice versa, decreasing risks led to higher judgments of benefits. However, we did not find any differences in the low-benefit conditions. Specifically, decreasing the benefits did not lead to higher judgments of risks (See Table S41–S44 in the supplementary material for detailed results).

In addition, we report two extensions. First, we found that the negative relationship between risks and benefits was stronger among participants who reported feeling higher incidental pleasure. Concurrently, people who felt pleasant may have generally relied more on heuristic processing—in this case the AH (Bohner et al., 1995). Previous findings, which manipulated negative mood, showed increased risk perceptions (Västfjäll et al., 2014). This may indicate that negative mood has a more pointed effect on risk-benefit judgments, although our findings cannot speak on this as we did not have a lot of data on the negative side of our measures, meaning we had few participants feeling low pleasure and low arousal (see Figure S5 in the supplementary material). This may have reduced our chances of obtaining more precise findings on how incidental affect can modulate the negative relationship. Furthermore, it is important to note that we measured naturally occurring incidental mood whereas previous research manipulated mood directly.

Second, we looked at whether manipulating risks or manipulating benefits impacts the strength of the negative relationship. Initially, our results showed the strength of the negative relationship was stronger when risks, as opposed to benefits, were manipulated. However, a more detailed look shows that this effect is most likely a product of the fact that there was no impact on the non-manipulated attribute in the low-benefit condition (see

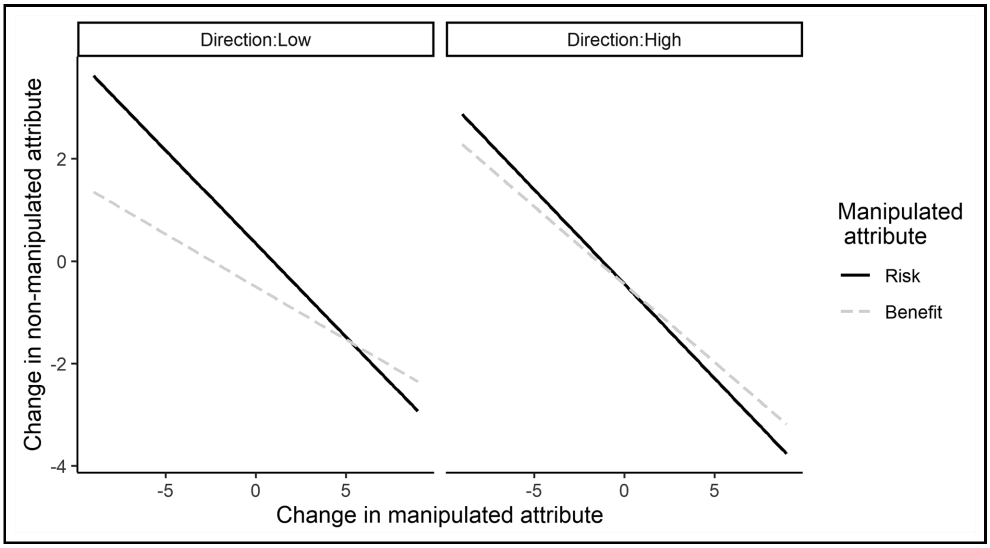

Figure 5. Relationship Between Manipulated and Non-Manipulated Attributes as a Function of Risk/Benefit Manipulations

Table S41–S44; the original findings seem to show this as well; see Table S34 and Table S35 in the supplementary material for detailed results). It is worth pointing out that manipulating low benefits did lead to a predicted change in benefits—people judged them as considerably lower (i.e., there was a successful manipulation; see Table S41–S44). But decreasing benefits did not lead to the predicted impact on risks. This may hint at the fact that providing low benefit info is not enough to lead to perceptible changes in affect as it may be that risks are simply better at evoking an affective reaction (cf. Pachur et al., 2014). Our results also hint at the fact that people may pay more attention to risks—both increase and decrease in risks—while this is not the case for benefits, where only increase in benefits led to perceivable changes. Alternatively, the lack of impact on the non-manipulated attribute in the low-benefit condition may hint at sensitivity to the actual relationship of risks and benefits in the world, namely, that they are often positively correlated. As mentioned in the introduction, technologies low in benefit are unlikely to be high in risk. It is of course not incommensurable that this sensitivity exists along a strong affective process that leads to negative relationships between risks and benefits.

Current findings may have important implications for risk communication (Thaler & Sunstein, 2008; Yang et al., 2014). For instance, communication efforts about new technologies ought to contend that risk information may overweigh other benefit information and is more malleable to manipulate. While out of scope for this research, it may be worth taking a closer look at what associations people might have with the terms "risky" and "beneficial." Specifically, people may already associate and interpret these terms as "bad" (for risky) and "good" (for beneficial), explaining the negative correlation.

We believe this replication strengthens the claim that it is possible to causally affect risk and benefit judgments. The negative relationship has been presented as a demonstration of the AH. However, while the effect is *consistent* with an AH, we (as the original finding) do not provide direct evidence that affect does mediate this negative relationship. Indeed, the negative relationship could also occur due to a more cognitive explanation. While we show evidence that change in the manipulated attribute is a mediator between the manipulations and non-manipulated attribute, this may be one of the potential mediators and the underlying cause remains uncovered. Some recent research has, for example, found more support for manipulations of availability by the recall, rather than affect, to have a stronger impact on how risk judgments are constructed (Efendić, 2021). Nevertheless, with this replication, we hope to encourage future researchers that this paradigm is robust and could potentially be used to tease apart any cognitive/affective explanations of risk/benefit judgments.

Finally, in our replication, we focused on the original three technological scenarios as the risky hazards. While one could argue that people's attitudes toward these risks have changed in the intervening 20 years since the original study, impacting the strength of the negative relationship, our results show similar effects. This could indicate that either the attitudes did not change, or, equally likely, that the manipulations of risk/benefit go well and beyond beliefs and attitudes. In that sense, future work should look at whether the negative relationship extends to other hazards. For instance, Skagerlund et al. (2020) found that the inverse relationship extends to numerous other hazards, activities, and technologies.

Author Contributions

G.F. led the project, supervised each step of the project, conducted the pre-registration, and ran data collection. E.E. and S.P.C. followed up on initial work by the other coauthors to verify and conduct additional analyses, and completed the manuscript draft. E.E., S.P.C., and G.F. jointly finalized the manuscript for submission. C.S.L., L.Y.Y., M.J.K., and C.Y.L. conducted the replication and extension as part of university course work. They conducted an initial analysis of the paper, designed the replication, initiated the extensions, wrote the pre-registration, conducted initial data analysis, and wrote initial replication reports.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research was supported by the Teaching Development Grant of the University of Hong Kong. S.P.C. thanks the Institute of International Business and Governance (IIBG), established with the substantial support of a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (UGC/IDS 16/17), for its support.

ORCID iD

Gilad Feldman https://orcid.org/0000-0003-2812-6599

Supplemental Material

The supplemental material is available in the online version of the article.

Notes

1. We would like to thank the original authors for providing the materials.
2. Note that the preregistrations follow a registered report format. This means that a manuscript-like document was produced reporting simulated random data results. Please see also the Read-me document in the wiki page on the OSF preregistrations here: https://osf.io/pg3ae/ for a detailed guide on where to find information on preregistered materials, design, and analysis plan.
3. Indicating a low proficiency of English, self-report not being serious about filling in the survey, who guessed the hypothesis, have done the survey before, who failed to complete the survey, and those not from the United States/ United Kingdom. Please see Table S2 in the supplement for more detail.
4. We ran all the models below with study included as a fixed effect and we did not find any evidence that the results differed between studies. Please see tables S9, S13, S18, and S23 in the supplement.
5. The additional experimental condition presented participants both information on risk *and* benefit. This presentation made it impossible to test the negative relationship and we saw fit to exclude it. Some 192 of the 968 participants in the prolific sample were in the excluded condition. Responses from remaining 776 prolific participants was included in the final analysis. Please see also note 2 in Table S1 and Table S5 in supplement for more details.
6. In the original study, the question added the phrasing ". . . to U.S. society as a whole" at the end. We used this exact phrasing in the MTurk sample (which included people from the United States) but decided to exclude this for the Prolific sample as these participants were from the United Kingdom.

References

Albarracín, D., & Kumkale, G. T. (2003). Affect as information in persuasion: A model of affect identification and discounting. *Journal of Personality and Social Psychology*, *84*(3), 453–469. https://doi.org/10.1037/0022-3514.84.3.453

Alhakami, A. S., & Slovic, P. (1994). A psychological study of the inverse relationship between perceived risk and perceived benefit. *Risk Analysis*, *14*(6), 1085–1096. https://doi.org/10.1111/j.1539-6924.1994.tb00080.x

Anderson, A. (1981). *Foundations of information integration theory*. Academic Press.

Andrade, E. B. (2005). Behavioral consequences of affect: Combining evaluative and regulatory mechanisms. *Journal of Consumer Research*, *32*(3), 355–362. https://doi.org/10.1086/497546

Bakker, M., van Dijk, A., & Wicherts, J. M. (2012). The rules of the game called psychological science. *Perspectives on Psychological Science*, *7*(6), 543–554. https://doi.org/10.1177/1745691612459060

Bateman, I., Dent, S., Peters, E., Slovic, P., & Starmer, C. (2007). The affect heuristic and the attractiveness of simple gambles. *Journal of Behavioral Decision Making*, *20*(4), 365–380. https://doi.org/10.1002/bdm.558

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Benthin, A., Slovic, P., Moran, P., Severson, H., Mertz, C. K., & Gerrard, M. (1995). Adolescent health-threatening and health-enhancing behaviors: A study of word association and imagery. *Journal of Adolescent Health*, *17*(3), 143–152. https://doi.org/10.1016/1054-139X(95)00111-5

Bhullar, N., Hine, D. W., Marks, A., Davies, C., Scott, J. G., & Phillips, W. (2014). The affect heuristic and public support for three types of wood smoke mitigation policies. *Air Quality, Atmosphere and Health*, *7*(3), 1–10. https://doi.org/10.1007/s11869-014-0243-1

Bohner, G., Moskowitz, G. B., & Chaiken, S. (1995). The interplay of heuristic and systematic processing of social information. *European Review of Social Psychology*, *6*(1), 33–68.

Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain*. Avon.

Efendić, E., Drače, S., & Ric, F. (2019). The combination of multiple affective experiences and their impact on valuation judgments. *Cognition and Emotion*, *34*(4), 684–699. https://doi.org/10.1080/02699931.2019.1675597

Efendic E. (2021). How do People Judge Risk? Availability may Upstage Affect in the Construction of Risk Judgments. *Risk analysis: an official publication of the Society for Risk Analysis*. Advance online publication. https://doi.org/10.1111/risa.13729

Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making*, *13*(1), 1–17. https://doi.org/10.1002/(SICI)1099-0771(200001/03)13:1<1::AID-BDM333>3.0.CO;2-S

Fischhoff, B., Slovic, P., Lichtenstein, S., Read, S., & Combs, B. (1978). How safe is safe enough? A psychometric study of attitudes towards technological risks and benefits. *Policy Sciences*, *9*(2), 127–152. https://doi.org/10.1007/BF00143739

Ganzach, Y. (2000). Judging risk and return of financial assets. *Organizational Behavior and Human Decision Processes*, *83*(2), 353–370. https://doi.org/10.1006/obhd.2000.2914

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*(1), 54–69. https://doi.org/10.1037/a0028347

Kahneman, D. (2003). A perspective on judgment and choice—Mapping bounded rationality. *American Psychologist*, *58*(9), 697–720. https://doi.org/10.1037/0003-066x.58.9.697

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus, and Giroux.

Keller, C., Siegrist, M., & Gutscher, H. (2006). The role of the affect and availability heuristics in risk communication. *Risk Analysis*, *26*(3), 631–639. https://doi.org/10.1111/j.1539-6924.2006.00773.x

King, J., & Slovic, P. (2014). The affect heuristic in early judgments of product innovations. *Journal of Consumer Behaviour*, *13*(6), 411–428. https://doi.org/10.1002/cb.1491

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(1), 1–26. https://doi.org/10.18637/jss.v082.i13

LeBel, E. P., McCarthy, R. J., Earp, B. D., Elson, M., & Vanpaemel, W. (2018). A unified framework to quantify the credibility of scientific findings. *Advances in Methods and Practices in Psychological Science*, *1*(3), 389–402. https://doi.org/10.1177/2515245918787489

LeBel, E. P., Vanpaemel, W., Cheung, I., & Campbell, L. (2019). A brief guide to evaluate replications. *Meta-Psychology*, *3*,1–9. https://doi.org/10.15626/MP.2018.843

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annual Review of Psychology*, *66*(1), 799–823. https://doi.org/10.1146/annurev-psych-010213-115043

Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*, *49*(2), 433–442. https://doi.org/10.3758/s13428-016-0727-z

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, *127*(2), 267–286. https://doi.org/10.1037/0033-2909.127.2.267

MacGregor, D. G., Slovic, P., Dreman, D., & Berry, M. (2000). Imagery, affect, and financial judgment. *The Journal of Psychology and Financial Markets*, *1*(2), 104–110. https://doi.org/10.1207/S15327760JPFM0102_2

Pachur, T., Hertwig, R., & Wolkewitz, R. (2014). The affect gap in risky choice: affect-rich outcomes attenuate attention to probability information. *Decision*, 1(1), 64.

Peters, E., Västfjäll, D., Gärling, T., & Slovic, P. (2006). Affect and decision making: A "hot" topic. *Journal of Behavioral Decision Making*, *19*(2), 79–85. https://doi.org/10.1002/bdm.528

Rachlin, H. (2003). *Bounded rationality: The adaptive toolbox* (Vol. 79). MIT Press.

Raue, M., D'Ambrosio, L. A., Ward, C., Lee, C., Jacquillat, C., & Coughlin, J. F. (2019). The influence of feelings while driving regular cars on the perception and acceptance of self-driving cars. *Risk Analysis*, *39*(2), 358–374. https://doi.org/10.1111/risa.13267

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145–172. https://doi.org/10.1037/0033-295X.110.1.145

Schwarz, N. (2012). Feelings-as-information theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 289–308). SAGE.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*(3), 513–523. https://doi.org/10.1037/0022-3514.45.3.513

Siegrist, M., & Sutterlin, B. (2014). Human and nature-caused hazards: The affect heuristic causes biased decisions. *Risk Analysis*, *34*(8), 1482–1494. https://doi.org/10.1111/risa.12179

Skagerlund, K., Forsblad, M., Slovic, P., & Västfjäll, D. (2019). The affect heuristic and risk perception—Stability across elicitation methods and individual cognitive abilities. *PsyArXiv* [Preprint]. https://doi.org/10.31234/osf.io/mpvu8

Skagerlund, K., Forsblad, M., Slovic, P., & Västfjäll, D. (2020). The affect heuristic and risk perception—Stability across elicitation methods and individual cognitive abilities. *Frontiers in Psychology, 11*, Article 970. https://doi.org/10.3389/fpsyg.2020.00970

Slovic, P. (1987). Perception of risk. *Science*, *236*(4799), 280–285. https://doi.org/10.1126/science.3563507

Slovic, P., Finucane, M., Peters, E., & MacGregor, D. G. (2002). Rational actors or rational fools: Implications of the effects heuristic for behavioral economics. *Journal of Socio-Economics*, *31*(4), 329–342. https://doi.org/10.1016/S1053-5357 (02)00174-9

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2007). The affect heuristic. *European Journal of Operational Research*, *177*(3), 1333–1352. https://doi.org/10.1016/j.ejor.2005.04.006

Slovic, P., & Västfjäll, D. (2010). Affect, moral intuition, and risk. *Psychological Inquiry*, *21*(4), 387–398. https://doi.org/10.1080/1047840X.2010.521119

Thaler, R., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.

Västfjäll, D., Peters, E., & Slovic, P. (2014). The affect heuristic, mortality salience, and risk: Domain-specific effects of a natural disaster on risk-benefit perception. *Scandinavian Journal of Psychology*, *55*(6), 527–532. https://doi.org/10.1111/sjop.12166

Vuorre, M., & Bolger, N. (2018). Within-subject mediation analysis for experimental data in cognitive psychology and neuroscience. *Behavior Research Methods*, *50*(5), 2125–2143. https://doi.org/10.3758/s13428-017-0980-9

Wyer, R. S., Clore, G. L., & Isbell, L. M. (1999). Affect and information processing. *Advances in Experimental Social Psychology*, *31*, 1–77. https://doi.org/10.1016/S0065-2601(08)60271-3

Yang, Z. J., Aloe, A. M., & Feeley, T. H. (2014). Risk information seeking and processing model: A meta-analysis. *Journal of Communication*, *64*(1), 20–41. https://doi.org/10.1111/jcom.12071

Author Biographies

Emir Efendić is a postdoctoral scholar at the School of Business and Economics in Maastricht University in the Netherlands. His research focuses on judgment and decision-making.

Subramanya Prasad Chandrashekar recently completed a research assistant professor position with the Lee Shau Kee School of Business and Administration at the Hong Kong Metropolitan University. His research focuses on social status, lay-beliefs, and judgment and decision-making.

Cheong Shing Lee is a student at the University of Hong Kong during the academic year 2019–2020.

Lok Yan Yeung is a student at the University of Hong Kong during the academic year 2019–2020.

Min Ji Kim is a student at the University of Hong Kong during the academic year 2019–2020.

Ching Yee Lee is a student at the University of Hong Kong during the academic year 2019–2020.

Gilad Feldman is an assistant professor with the University of Hong Kong psychology department. His research focuses on judgment and decision-making.

Handling Editor: Lissa Libby