Received: 2 April 2021 | Revised: 9 February 2022 | Accepted: 11 February 2022

DOI: 10.1002/bdm.2278

RESEARCH ARTICLE

WILEY

How false feedback influences decision-makers' risk preferences

Petko Kusev[1] | Paul van Schaik[2] | Joseph Teal[1] | Rose Martin[3] | Lars Hall[4] | Petter Johansson[4]

[1]Behavioural Research Centre, Huddersfield Business School, University of Huddersfield, Huddersfield, UK

[2]Department of Psychology, Teesside University, Middlesbrough, UK

[3]Department of People and Organisations, Surrey Business School, University of Surrey, Guildford, UK

[4]Lund University Cognitive Science, Lund University, Lund, Sweden

Correspondence
Petko Kusev, Behavioural Research Centre, Huddersfield Business School, University of Huddersfield, Huddersfield HD1 3DH, UK.
Email: p.kusev@hud.ac.uk

Funding information
British Academy, Grant/Award Number: SG 47881; Japan Society for the Promotion of Science; The Swedish Research Council, Grant/Award Number: 2014-1371

Abstract

Recent decision-making research provides empirical evidence that human risk preferences are constructed "on the fly" during risk elicitation, influenced by the decision-making context and the method of risk elicitation (Kusev et al., 2020). In this article, we explore the lability of human risk preferences and argue that the most recent choices guide decision-making. Accordingly, our novel proposal and experimental method provide a psychological tool that measures people's shift in preferences. Specifically, in our experiment (240 participants, registered UK users of an online survey panel), we developed and employed a two-stage risk elicitation experimental method. The results from the experiment revealed that providing participants with false feedback on their initial decisions (stage 1) changes their risk preferences at the feedback stage of the experiment in the direction of the false feedback. Moreover, participants' final decisions (stage 2) were influenced by the type of feedback (correct or false) and informed by their altered risk preferences at the feedback stage of the experiment. In conclusion, our work provides experimental evidence that human preferences are constructed "on the fly," influenced by the decision-making context and recent decision-making experience (e.g., Kusev et al., 2020; Slovic, 1995).

KEYWORDS

decision experience, decision-making under risk, false feedback, preference reversals, prospect theory

1 | INTRODUCTION

The foundation of economic normative theory is the idea that people have stable and identifiable preferences that inform their decisions (Elster, 1986). While some descriptive psychological theories of decision-making (e.g., Kahneman & Tversky, 1979; Tversky & Kahneman, 1992) can account for many of the phenomena not anticipated by normative theory (von Neumann & Morgenstern, 1947), most of these theories predict consistent (normatively rational or irrational) decision-making preferences (cf. Brandstätter et al., 2006; Hertwig et al., 2004; Kusev et al., 2009, 2020). For example, the fourfold pattern of risk preferences (risk aversion for low-probability loss and high-probability gain and risk-seeking for low-probability gain and high-probability loss) provides empirical evidence for the predictions by prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). In contrast, in this article, we explore the lability of (normatively rational or irrational) human preferences in risky decision-making scenarios. We propose that participants' most recent choices guide their subsequent risk preferences. Specifically, we expect that providing participants with false feedback on their initial decisions changes their risk preferences at the feedback stage of the experiment in the direction of the false feedback. Moreover, we also anticipate that the effect of the feedback on participants' final decisions is mediated by their altered risk preferences at the feedback stage of the experiment.

1.1 | The lability of preference

In contrast to normative and descriptive behavioral predictions, empirical evidence from decision-making research indicates that human risk preferences are constructed "on the fly," influenced by the decision-making context and method for preference elicitation (e.g., Kusev et al., 2009, 2020; Lichtenstein & Slovic, 1971; Slovic, 1995; Stewart et al., 2003, 2006; Zhang & Slovic, 2018). For example, developing further the certainty equivalent experimental method (originally proposed by Tversky & Kahneman, 1992), Kusev et al. (2020) found that decision-makers' risk preferences (normatively rational or irrational) are constructed during risk elicitation "on the fly." Specifically, variation in the decision context (linear or logarithmic scaling of the sure options around the expected value of certain options) induces preference reversals during risk elicitation and determines people's preferences in the domains of loss and gain. Moreover, the lability of human risk preferences is also evident in the result of weak association between self-reported and behavioral measures of risk preferences found by Frey et al. (2017).

Experimental evidence has also demonstrated that the memory of everyday experiences of risky events can also influence risky preferences (e.g., Kusev et al., 2009, 2017; Kusev & Van Schaik, 2011; Sekścińska et al., 2021; Stewart et al., 2003; Vlaev et al., 2010; Weiss-Cohen et al., 2021). For example, prospects that are presented in an insurance context are judged with greater risk aversion than mathematically identical choices presented as standard gambles (see Kusev et al., 2009). Accordingly, the authors concluded that the exaggerated risk in precautionary decisions is caused by the accessibility of risky events in memory; thus, participants' memory of previous experiences "leak" into subsequent precautionary choices. This suggests that despite risk information being explicitly provided, people's previous experiences of events influence their decisions.

More generally, behavioral science researchers have demonstrated that memories need not even be real to influence people's behavior. Specifically, in the typical false feedback paradigm, participants are misled about the occurrence of a past event, so that the impact of the false memory on their subsequent behavior can be examined (Pezdek & Freyd, 2009). For example, Geraerts et al. (2008) demonstrated that participants were less likely to eat an egg salad if it had been falsely suggested to them that, as a child, they had become ill after eating an egg salad. Nevertheless, this manipulation does not extend to foods which people enjoy and frequently consume (Pezdek & Freyd, 2009). However, people can also be "implanted" with more far-fetched false memories (e.g., getting lost in a shopping mall; Loftus & Pickrell, 1995). Accordingly, false information and fabricated evidence can be used to elicit false memories and inaccurate eyewitness testimonies and induce people to testify about crimes which they never witnessed (see Wade et al., 2010). In one particular study, Loftus (1975) asked participants to watch footage of a multiple-car accident and then, having watched the film, to answer a series of questions about the accident. The results revealed that participants asked questions starting with "did you see the" (e.g., "did you see the broken headlight?") were more likely to report having seen an object (e.g., a broken headlight) than participants asked questions starting with "did you see a" (e.g., "did you see a broken headlight?"). Moreover, a related line of research employing the choice blindness methodology has revealed that false feedback can influence non-risky preferences and behavior (e.g., Hall et al., 2010, 2012; Johansson et al., 2005, 2006, 2008; Lind et al., 2014; Steenfeldt-Kristensen & Thornton, 2013). However, in the choice blindness method, the likelihood that participants will detect false feedback is determined by their overall preference for the choice options (Somerville & McGowan, 2016). In particular, participants are more likely to detect the false feedback manipulation if they have a strong overall preference for a particular decision option (e.g., a particular type of chocolate).

Moreover, false feedback can also influence learning. For instance, in a recent experiment, Wang et al. (2019) demonstrated that false feedback can alter the strength of people's memory associations between stimuli and influence reinforcement decision-making. Specifically, in a task where participants learnt associations between neutral stimuli (images and patterned circles) and rewards, those given false feedback were less likely to choose stimuli associated with a reward than those given correct feedback. Furthermore, participants given false rather than correct feedback reported decreased recollection ratings for their memory associations. In other words, the strength of participants' memory associations was influenced (decreased) by false feedback.

However, despite the broad exploration and use of false feedback (and false memory) methods in psychology, no previous research has explored the influence of false feedback on human risky preferences, particularly in relation to predictions of decision-making theories and their methods. Accordingly, it is plausible that participants' most recent choices guide their subsequent risk preferences (see also Ariely & Norton, 2008, for a review). Thus, in this research, we explore this possibility.

In studies of risk and decision-making, a ubiquitous tool for both economists and psychologists is the monetary-gamble method (e.g., Birnbaum, 2008; Gonzalez & Wu, 1999; Lopes, 1983; Prelec, 1998; Tversky & Kahneman, 1992). With the advent of neuro-economics as a separate and high-profile subfield of economics and cognitive science, the use of monetary gambles has become even more prominent (e.g., Camerer et al., 2005; Glimcher & Rustichini, 2004; Kenning & Plassmann, 2005). In these studies, it is commonly assumed that the behavioral patterns uncovered in experiments using monetary gambles can be used as a method of preference elicitation for decisions of all kinds.

In response to these opportunities, we see the monetary (gamble) decision-making methods employed by theorists from psychology, economics, and experimental philosophy as an important domain in which to test our novel proposal. For example, is it possible that the participants will accept the substituted risky choices as their own (via false feedback), even when the decision alternatives are so explicitly and quantitatively specified? This question also opens up the next step of inquiry: Will the participants come to prefer the substituted risky choices when the choices are made again; what would happen to their risk preferences over time? Answers to these questions will provide fundamental insights to the field of decision-making under risk, especially since risk preferences are supposed to be dispositions that remain stable over time according to most psychological theories of decision-making.

In order to investigate and further support our claim about the fragile nature of human risk preferences, we developed and tested a new experimental method that includes correct or false feedback (see Figure 1), based on the certainty equivalent experimental task proposed by Kusev et al. (2020). In this experimental task, participants' preferences are explored in the domains of loss and gain. Specifically, participants make a series of choices between sure and probabilistic gambles (in the domains of loss or gain). After the participants had completed their choices (decision stage 1 of the experiment), they are presented with all the gamble prospects again with their previous choices highlighted (providing correct or false feedback) and asked to confirm their choices from stage 1 by choosing again (feedback stage of the experiment). For example, in the false feedback gain or loss conditions (feedback stage of the experiment), when the participants are asked to confirm their choices, the opposite of their original choice was highlighted, thereby giving false feedback (see Figure 1). The gamble prospects are then presented a third time, and participants make their choices once again (decision stage 2 of the experiment).

According to normative and descriptive predictions, decision preferences (normatively rational or irrational) are stable. However, by employing this method, we found that respondents' risk preferences shifted in the direction of the false feedback. To summarize our argument, we propose the following hypotheses:

H1. False feedback changes respondents' risk preferences in the direction of the false feedback.

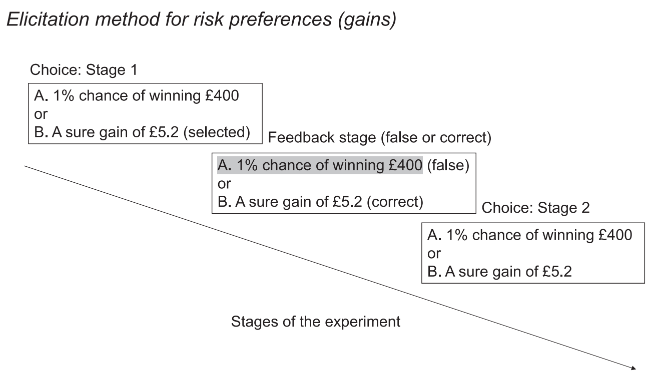

FIGURE 1 Elicitation method for risk preferences (gains)

H2. Risk preferences immediately after feedback mediate the effect of feedback on subsequent risk preferences.

2 | EXPERIMENT

2.1 | Participants

Participants were 240 (139 female and 101 male) registered UK users of an online survey panel who completed the experiment. Mean age was 41 years ($SD = 13.19$). Participants took part individually, received a payment of £1, and spent approximately 17 min to complete the experiment (no further time measures were taken); 16 participants did not complete the decision trials and were excluded from the study (240 participants in total, excluding those 16). The experiment received departmental research ethics committee approval (Department of Psychology, Kingston University London). All participants were randomly assigned to the experimental conditions and treated in accordance with the ethical standards of the British Psychological Society and APA ethical principles; due to the false feedback manipulation, participants were fully debriefed about the goal of the experiment and were given the opportunity to withdraw their responses.

For statistical testing, we used a significance level of .05. Although we did not assume an effect size, we wanted to ensure that our sample size would allow us to detect a large effect size ($f = .25$ by convention; Cohen, 1988). We ran the experiment for 12 days to ensure that data collection from a sufficiently large sample would achieve a statistical power of at least .95. According to the retrospective power analysis, the achieved sample size ($N = 240$) produced a power of ≥.95 for all effects in 2 × 2 × (2) analysis of variance (ANOVA) and 2 × 2 analysis of covariance (ANCOVA), which was sufficient to achieve our target. The data that support the findings of this study are available from the corresponding author upon request.

2.2 | Experimental design and procedure

A mixed-measures 2 × 2 × (2) design was used, with the independent between-subject variables *domain of decision-making* (gain or loss) and *feedback* (correct or false), and independent within-subject variable *stage of the experiment* (stage 1 and feedback stage). The dependent variable was respondents' risk preferences (proportion of risk-averse preferences; between 0: risk-seeking and 1: risk aversion). In addition, an independent-measures 2 × 2 design was used, with *domain of decision-making* (gain or loss) and *feedback* (correct or false), with participants' risk preferences at stage 1 as the covariate, the same dependent variable at stage 2. Finally, mediation analyses were conducted to test whether the respondents' altered risk preferences at the feedback stage of the experiment are a mediator of the relationship between type of feedback (correct or false) at the feedback stage and risk preferences at stage 2 of the experiment.

At the beginning of the study, task instructions, an example scenario with illustrative choices, and then binary decision-making tasks were presented to all participants in an online computer-based experiment. For example, participants were presented with tasks to choose between a risky option A and certain option B (by clicking on the preferred option):

Option A. 1% chance of winning £400, or
Option B. A sure gain of £5.2.

All respondents were presented with and completed 189 trials of binary decision-making scenarios (between a probabilistic and certain/sure options): 63 in stage 1, 63 in feedback stage (correct or false), and 63 in stage 2 of the decision-making experiment (see Figure 1).

All participants completed 63 choices in stage 1 of the experiment; next, they were presented with all the prospects again with their previous choices highlighted (providing false [the opposite of participants' original choice] or correct feedback [the same as participants' original choice]). In this feedback stage of the experiment, the respondents were asked to confirm their choices from stage 1 by choosing again. Accordingly, we anticipated that by providing false feedback on the decisions made in stage 1, in the feedback stage of the experiment, participants will shift their preferences in the direction of the false feedback. Therefore, we expected reduced risk aversion in the domain of gain and reduced risk-seeking in the domain of loss. In stage 2 of the experiment, the gamble prospects were presented a third time, and the same participants made their choices once again (see Figure 1). In each stage, all decision trials were randomized.

2.3 | Decision stimuli

An interactive online computer-based task for binary decision-making was developed (with Qualtrics XM) and used. We developed and employed a two-stage certainty equivalent risk elicitation method where participants make choices in binary decision-making scenarios (choices between probabilistic and sure options). Following on from our previous work (Kusev et al., 2020), we used linearly spaced sure options (equal number of risk-averse and risk-seeking sure options balanced around the expected value EV for each probability level) to present the binary-choice prospects. For example, for risky prospects with 1% chance of winning £400 (EV = £4), there were three sure options above the EV and three sure options below the EV: £0.4, £1.6, £2.8, £4, £5.2, £6.4, and £7.6 (incremental and decremental steps of £1.2).

Four types of binary decision-making scenario were included in the feedback stage of the experiment, each corresponding with one of the following experimental conditions: gamble-gain tasks with false feedback, gamble-gain tasks with correct feedback, gamble-loss tasks with false feedback, or gamble-loss tasks with correct feedback. In stage 1 and stage 2, the following scenarios were presented: gamble gain or gamble loss without feedback. Accordingly, the decision trials were generated by

i. combining a monetary amount (£400—probabilistic outcome) with nine probability levels (.01, .05, .10, .25, .50, .75, .90, .95, and .99); hence, nine probabilistic combinations were presented with

TABLE 1 Range of sure options linearly spaced around the expected value of probabilistic options' expected values

p (win £400)	Expected value	Min sure amount	Max sure amount
.01	£4	£0.40	£7.60
.05	£20	£16.40	£23.60
.10	£40	£36.40	£43.60
.25	£100	£96.40	£103.60
.50	£200	£196.40	£203.60
.75	£300	£296.40	£303.60
.90	£360	£356.40	£363.60
.95	£380	£376.40	£383.60
.99	£396	£392.40	£399.60
***p* (lose £400)**	**Expected value**	**Min sure amount**	**Max sure amount**
.01	−£4	−£0.40	−£7.60
.05	−£20	−£16.40	−£23.60
.10	−£40	−£36.40	−£43.60
.25	−£100	−£96.40	−£103.60
.50	−£200	−£196.40	−£203.60
.75	−£300	−£296.40	−£303.60
.90	−£360	−£356.40	−£363.60
.95	−£380	−£376.40	−£383.60
.99	−£396	−£392.40	−£399.60

ii. one of the seven sure monetary amounts (linearly spaced sure options balanced around the EV for each probability level [three above and three below the EV]). For example, for 1% chance winning £400, there were seven linearly spaced sure options (£0.4, £1.6, £2.8, £4, £5.2, £6.4, and £7.6) (see Table 1). Therefore, there were 9 (probabilities) × 7 (linearly spaced sure options) × 3(stage 1, feedback stage, and stage 2 of the experiment) = 189 decision trials.

Following the certainty equivalent method, participants' risk preferences are computed from the sure options. Specifically, the certainty equivalent is the midpoint between the lowest accepted sure value and the highest rejected sure value (for each of the probability levels); certainty equivalent values above the expected value indicate risk-seeking preferences, while certainty equivalent values below the expected value indicate risk-averse preferences (also see Kusev et al., 2020). Accordingly, we have measured and established participants' risk preferences for each of the probability levels, in stage 1, the feedback stage, and stage 2 of the experiment. Thus, the dependent variable was respondents' risk preferences as proportion of risk-averse preferences (between 0: risk-seeking and 1: risk aversion).

2.4 | Results

2.4.1 | Risk preferences in stage 1 and feedback stage (manipulation check)

The aim of the false feedback was to change respondents' risk preferences in the direction of the feedback (H1). Therefore, the change in respondents' risk preferences between stage 1 and feedback stage of the experiment was tested. A mixed-measures 2 × 2 × (2) ANOVA was conducted, with the independent variables *domain of decision-making* (gain or loss) and *feedback* (correct or false), within-subject variable *stage of the experiment* (stage 1 and feedback stage), and the dependent variable risk preferences.

The results revealed that the main effects of feedback $F < 1$, stage of the experiment, $F(1, 236) = 1.78$, $p = .183$, $\eta_p^2 = .01$, as well as the two-way interaction feedback by stage of the experiment, $F(1, 236) = 1.35$, $p < .246$, $\eta_p^2 = .01$, were not statistically significant. However, the main effect of domain of decision-making, $F(1, 236) = 322.70$, $p < .001$, $\eta_p^2 = .58$, as well as the two-way interactions domain of decision-making by feedback, $F(1, 236) = 17.52$, $p < .001$, $\eta_p^2 = .07$, domain of decision-making by stage of the experiment, $F(1, 236) = 35.42$, $p < .001$, $\eta_p^2 = .13$, and the three-way interaction domain of decision-making by feedback by stage of the experiment, $F(1, 236) = 27.10$, $p < .001$, $\eta_p^2 = .10$, were all significant (see Figure 2).

As the three-way interaction was significant, interpretation of the two-way interactions and main effects was precluded. Accordingly, four follow-up simple-effect tests by domain of decision-making (gain or loss) and feedback (correct or incorrect) were conducted.

With correct feedback, the differences between participants' risk preferences in stage 1 and feedback stage of the experiment, in the domains of gain $t(59) = 0.77$, $p = .447$, $d = 0.10$, and loss $t(59) = -1.88$, $p = .065$, $d = -0.24$, were not significant (see Figure 2). However, with false feedback in the domains of gain and loss, participants changed their preferences from stage 1 in the direction of the false feedback manipulation (feedback stage of the experiment). Specifically, in the domain of gain, participants were significantly less risk-averse in the feedback stage of the experiment ($M = 0.68$, $SD = 0.34$) than in stage 1 of the experiment ($M = 0.85$, $SD = 0.23$), $t(59) = 3.58$, $p = .001$, $d = 0.46$ (see Figure 2). In the domain of loss, participants were significantly less risk-seeking in the feedback stage of the experiment ($M = 0.51$, $SD = 0.25$) than stage 1 of the experiment ($M = 0.24$, $SD = 0.30$), $t(59) = -4.47$, $p < .001$, $d = 0.58$ (see Figure 2).

Therefore, the manipulation check was successful: With false feedback, participants' risk preferences changed significantly from stage 1 to the feedback stage of the experiment in the direction of the false feedback manipulation.

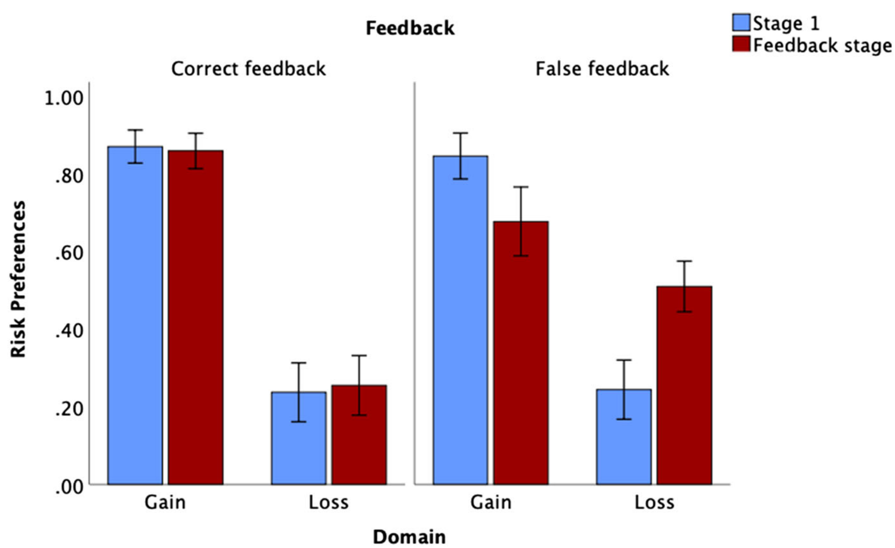

FIGURE 2 Risk preferences by domain of decision-making, type of feedback, and stage of the experiment; mean values with error bars (95% confidence intervals of the mean) represent the proportion of participants' risk-averse preferences (between 0: risk-seeking and 1: risk aversion) [Colour figure can be viewed at wileyonlinelibrary.com]

2.4.2 | Risk preferences in stage 2 of the experiment

In order to test H1 at stage 2 of the experiment, it was necessary to hold initial risk preferences from stage 1 constant. Therefore, an independent-measures 2 × 2 ANCOVA was conducted on the participants' risk preferences at stage 2 of the experiment, with participants' risk preferences at stage 1 as the covariate, and domain of decision-making (gain or loss) and feedback (correct or false) as the independent variables.

The results revealed that participants' risk preferences at stage 1 were strongly associated with their risk preferences at stage 2 of the experiment, $F(1, 235) = 55.78$, $p < .001$, $\eta_p^2 = .19$. Moreover, the effect of domain of decision-making, $F(1, 235) = 15.02$, $p < .001$, $\eta_p^2 = .06$, as well the two-way interaction domain of decision-making by feedback, $F(1, 235) = 19.74$, $p < .001$, $\eta_p^2 = .08$, was significant. However, the main effect of feedback was statistically not significant, $F < 1$.

Because the two-way interaction was significant, follow-up analysis of covariance by domain of decision-making (gain or loss) was conducted. The results showed that in the domain of gain, participants' risk preferences at stage 1 were strongly associated with their risk preferences at stage 2 of the experiment, $F(1, 117) = 26.38$, $p < .001$, $\eta_p^2 = .18$, and that the effect of feedback was significant, $F(1, 117) = 8.54$, $p = .004$, $\eta_p^2 = .07$, with reduced risk aversion when the feedback was false (see Figure 3). Specifically, the participants were significantly less risk-averse in the false feedback condition ($M = 0.73$, $SD = 0.32$) than in the correct feedback condition of the experiment ($M = 0.87$, $SD = 0.18$). Furthermore, the results showed that in the domain of loss, participants' risk preferences at stage 1 were strongly associated with their risk preferences at stage 2 of the experiment, $F(1, 117) = 30.17$, $p < .001$, $\eta_p^2 = .21$, and that the effect of feedback was significant, $F(1, 117) = 10.94$, $p = .001$, $\eta_p^2 = .09$, with reduced risk-seeking when the feedback was false (see Figure 3). Specifically, the participants were significantly less risk-seeking in the false feedback condition ($M = 0.40$, $SD = 0.26$) than in the correct feedback condition of the experiment ($M = 0.24$, $SD = 0.29$).

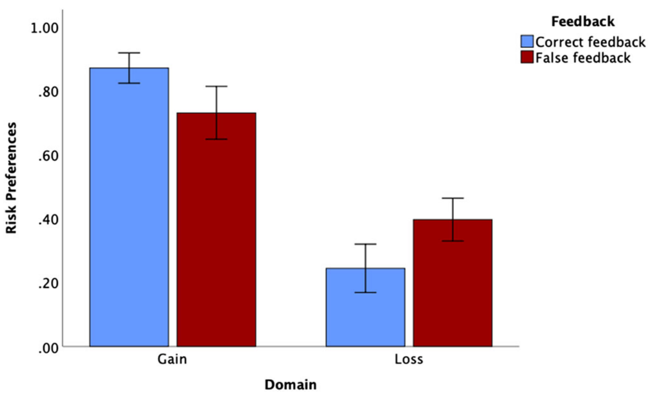

FIGURE 3 Risk preferences by domain of decision-making and type of feedback at stage 2; mean values with error bars (95% confidence intervals of the mean) represent the proportion of participants' risk-averse preferences (between 0: risk-seeking and 1: risk aversion) [Colour figure can be viewed at wileyonlinelibrary.com]

Providing false feedback to the participants influenced their risk preferences in the direction of the false feedback manipulation, even when we held participants' risk preferences from stage 1 constant by analyzing these as a covariate.

2.4.3 | Predicting risk preferences (stage 2)

In order to test H2, two mediation analyses (by domain of decision-making: model A [gain] or model B [loss]) were conducted to test whether the respondents' altered risk preferences at the feedback stage of the experiment are a mediator of the relationship between type of feedback (correct or false) and risk preferences at stage 2 of the experiment. The predictor variable was type of feedback, the mediator was respondents' risk preferences at the feedback stage of the experiment, and the outcome variable was respondents' risk preferences at stage 2 of the experiment. The indirect effect of feedback through the mediator risk preferences at the feedback stage of the experiment on risk preferences at stage 2 was tested by bootstrapping with $N = 1000$ (Hayes, 2018).

The results showed that participants' risk preferences at the feedback stage of the experiment were a mediator of the relationship between feedback (correct or incorrect) and participants' risk preferences at stage 2 of the experiment (see Table 2). Specifically, the results revealed that in the domain of gain (model A), the standardized indirect effect of feedback through the mediator participants' risk preferences at the feedback stage of the experiment was significant and negative (see Table 2); the participants were less risk-averse in the false feedback condition than in the correct feedback condition. Moreover, in the domain of loss (model B), the standardized indirect effect of feedback through the mediator participants' risk preferences at the feedback stage of the experiment was significant and positive (see Table 2); the participants were less risk-seeking in the false feedback condition than in the correct feedback condition. Accordingly, participants' final decisions (stage 2) were influenced by the type of feedback (correct or false) and informed by their altered risk preferences at the feedback stage of the experiment.

3 | DISCUSSION

For decades, the leading theories in decision-making research have informed us about how rational human agents should behave (normative decision theory; von Neumann & Morgenstern, 1947) and how humans actually behave (descriptive decision theory; Tversky & Kahneman, 1992). These theoretical accounts of human decision-making differ in their approach to understanding human decision-making processes and behaviors (normative vs. descriptive), yet both share a common assumption: Decision-making preferences are both stable and coherent. In contrast to this claim for decision stability, and in line with previous research (see Kusev et al., 2009, 2020;

TABLE 2 Mediation analysis by domain of decision-making (model A [gain] and model B [loss])

				Total effect			Direct effect				95% CI (BCa)	
Model	$F(2, 117)$	p	R^2	β	t	p	β	t	p	Indirect effect	LL	UL
A	296.95	<.001	.84	−.14	−2.94	.004	.02	0.83	.410	−.16	−.245	−.074
B	142.67	<.001	.71	.15	3.02	.003	−.06	−1.88	.063	.21	.132	.300

Note: Values for total effect, direct effect, and indirect effect are standardized.

Lichtenstein & Slovic, 1971; Shafir et al., 1993; Slovic, 1995; Stewart et al., 2003, 2006), our novel proposal and experimental method aim to counter the argument for preference consistency and instead provide unique demonstrations of a shift in preferences (by means of providing false feedback on previous decisions).

Accordingly, we explored the lability of (normatively rational or irrational) human preferences in risky decision-making scenarios and tested our prediction that participants' most recent choices guide their subsequent risk preferences. In our novel approach, we developed and employed a two-stage risk elicitation experimental method where participants make choices in binary decision-making scenarios.

The results from the experiment revealed a shift in respondents' decision-making preferences for binary-choice prospects with false feedback in the domains of gain and loss. Specifically, when false feedback was administered, participants changed their preferences from stage 1 in the direction of the false feedback manipulation at the feedback stage of the experiment. In the domain of gain, participants were significantly less risk-averse in the feedback stage of the experiment than in stage 1 of the experiment. However, in the domain of loss, participants were significantly less risk-seeking in the feedback stage of the experiment than stage 1 of the experiment.

Our results also show that type of feedback significantly influenced participants' final decisions (stage 2), even when we held participants' risk preferences from stage 1 constant. In the domain of gain, the participants were significantly less risk-averse in the false feedback condition than in the correct feedback condition of the experiment in stage 2. In the domain of loss, the participants were significantly less risk-seeking in the false feedback condition than in the correct feedback condition of the experiment. Accordingly, providing false feedback to the participants influenced their risk preferences in the direction of the false feedback manipulation. Crucially, these findings were consistent with empirical findings from Kusev et al. (2009), where memory of previous experiences "leak" into (and therefore influence) subsequent decision-making.

We also found that decision-makers' risk preferences at the feedback stage of the experiment are a mediator of the relationship between type of feedback (correct or false) and respondents' risk preferences at stage 2 of the experiment. Specifically, as we predicted, participants' final decisions (stage 2) were influenced by the type of feedback (correct or false) and informed by their altered risk preferences at the feedback stage of the experiment.

Given our results, that providing participants with false feedback on their initial decisions changes their subsequent risk preferences, future research should explore whether a smaller or larger number of decision trials with false feedback would influence preferences in the same way as we found in our research. It is plausible that with fewer decision trials, providing false feedback to participants will surprise them and thus have a high salience. It is also worth exploring whether expanding the range of stimuli to include gambles with a wider range of amounts as well as incentivization of participants would change the labile risk preferences (once again).

This question has been explored with non-risky choices. For instance, in a choice blindness task with non-risky stimuli (chocolate), Somerville and McGowan (2016) found that incentivization plays a minimal role in the detection of false feedback. An additional question is whether participants' detection of false feedback will improve when they have strong choice preferences. Indeed, Somerville and McGowan (2016) found that when children had a strong preference for a particular product (chocolate), choice blindness was significantly less prevalent than with choice between female faces. Accordingly, with chocolate (but not faces), the children were likely to detect manipulations of their choices.

The results from the study reveal respondents' shift in risk preferences and support our prediction that participants' most recent choices guide their subsequent risk preferences. Thereby, our work further supports existing evidence from decision-making research that preferences are constructed "on the fly," influenced by the decision-making context and recent decision-making experience (e.g., Kusev et al., 2009, 2017, 2020; Lichtenstein & Slovic, 1971; Slovic, 1995; Stewart et al., 2006; Teal et al., 2021). However, it is important to note that the theoretical proposal for adaptive risk preferences that are constructed "on the fly" (e.g., Kusev et al., 2020) might be evolutionary supported (and beneficial). For example, human cognition and behaviors, crucial for survival, are based upon biologically and psychologically motivated learning and adaptation processes (e.g., Confer et al., 2010; Fawcett et al., 2014; Santos & Rosati, 2015; Simon, 1990). Accordingly, future research should explore what the benefit of constantly updating our preferences "on the fly" is.

ACKNOWLEDGMENTS

This article and the first author were supported in part by the Japan Society for the Promotion of Science (Tokyo) and the British Academy SG 47881. We also thank the researchers working in Kazuo Shigemasu's lab (University of Tokyo) and Kimihiko Yamagishi's lab (Tokyo Institute of Technology) for their comments and suggestions. Petter Johansson and Lars Hall wish to thank The Swedish Research Council (2014-1371).

ORCID

Petko Kusev https://orcid.org/0000-0002-0308-927X

REFERENCES

Ariely, D., & Norton, M. I. (2008). How actions create—not just reveal—preferences. *Trends in Cognitive Sciences: Neuroeconomics*, *12*(1), 13–16.

Birnbaum, M. H. (2008). New paradoxes of risky decision making. *Psychological Review*, *115*(2), 463–501.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review*, *113*(2), 409–432.

Camerer, C., Loewenstein, G., & Prelec, D. (2005). Neuroeconomics: How neuroscience can inform economics. *Journal of Economic Literature*, *43*(1), 9–64.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. Erlbaum.

Confer, J. C., Easton, J. A., Fleischman, D. S., Goetz, C. D., Lewis, D. M. G., Perilloux, C., & Buss, D. M. (2010). Evolutionary psychology: Controversies, questions, prospects, and limitations. *The American Psychologist*, *65*(2), 110–126. https://doi.org/10.1037/a0018413

Elster, J. (Ed.) (1986). *Rational choice*. New York University Press.

Fawcett, T. W., Fallenstein, B., Higginson, A. D., Houston, A. I., Mallpress, D. E. W., Trimmer, P. C., & McNamara, J. M. (2014). The evolution of decision rules in complex environments. *Trends in Cognitive Sciences*, *18*(3), 153–161. https://doi.org/10.1016/j.tics.2013.12.012

Frey, R., Pedroni, A., Mata, R., Rieskamp, J., & Hertwig, R. (2017). Risk preference shares the psychometric structure of major psychological traits. *Science Advances*, *3*(10), e1701381.

Geraerts, E. G., Bernstein, D., Merckelbach, H. L. G. J., Linders, C., Raymaekers, L. H. C., & Loftus, E. F. (2008). Lasting false beliefs and their behavioral consequences. *Psychological Science*, *19*(8), 749–753. https://doi.org/10.1111/j.1467-9280.2008.02151.x

Glimcher, P. W., & Rustichini, A. (2004). Neuroeconomics: The consilience of brain and decision. *Science*, *306*(5695), 447–452.

Gonzalez, R., & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology*, *38*(1), 129–166.

Hall, L., Johansson, P., & Strandberg, T. (2012). Lifting the veil of morality: Choice blindness and attitude reversals on a self-transforming survey. *PLoS ONE*, *7*(9), e45457.

Hall, L., Johansson, P., Tärning, S., Sikström, S., & Deutgen, T. (2010). Magic at the marketplace: Choice blindness for the taste of jam and the smell of tea. *Cognition*, *117*(1), 54–61.

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). Guilford Press.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, *15*(8), 534–539.

Johansson, P., Hall, L., & Sikström, S. (2008). From change blindness to choice blindness. *Psychologia*, *51*(2), 142–155.

Johansson, P., Hall, L., Sikström, S., Tärning, B., & Lind, A. (2006). How something can be said about telling more than we can know: On choice blindness and introspection. *Consciousness and Cognition*, *15*(4), 673–692.

Johansson, P., Sikström, S., & Olsson, A. (2005). Failure to detect mismatches between intention and outcome in a simple decision task. *Science*, *310*(5745), 116–119.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–291.

Kenning, P., & Plassmann, H. (2005). NeuroEconomics: An overview from an economic perspective. *Brain Research Bulletin*, *67*(5), 343–354.

Kusev, P., Purser, H., Heilman, R., Cooke, A. J., Van Schaik, P., Baranova, V., Martin, R., & Ayton, P. (2017). Understanding risky behavior: The influence of cognitive, emotional and hormonal factors on decision-making under risk. *Frontiers in Psychology*, *8*, 102. https://doi.org/10.3389/fpsyg.2017.00102

Kusev, P., & Van Schaik, P. (2011). Preferences under risk: Content-dependent behavior and psychological processing. *Frontiers in Psychology*, *2*, 269. https://doi.org/10.3389/fpsyg.2011.00269

Kusev, P., van Schaik, P., Ayton, P., Dent, J., & Chater, N. (2009). Exaggerated risk: Prospect theory and probability weighting in risky choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(6), 1487–1505.

Kusev, P., van Schaik, P., Martin, R., Hall, L., & Johansson, P. (2020). Preference reversals during risk elicitation. *Journal of Experimental Psychology: General*, *149*(3), 585–589.

Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology*, *89*(1), 46–55.

Lind, A., Hall, L., Breidegard, B., Balkenius, C., & Johansson, P. (2014). Speakers' acceptance of real-time speech exchange indicates that we use auditory feedback to specify the meaning of what we say. *Psychological Science*, *25*(6), 1198–1205.

Loftus, E. F. (1975). Reconstructing memory: The incredible eyewitness. *Jurimetrics Journal*, *15*(3), 188–193.

Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals*, *25*(12), 720–725. https://doi.org/10.3928/0048-5713-19951201-07

Lopes, L. L. (1983). Some thoughts on the psychological concept of risk. *Journal of Experimental Psychology. Human Perception and Performance*, *9*(1), 137–144. https://doi.org/10.1037/0096-1523.9.1.137

Pezdek, K., & Freyd, J. J. (2009). The fallacy of generalizing from egg salad in false-belief research. *Analyses of Social Issues and Public Policy*, *9*(1), 177–183. https://doi.org/10.1111/j.1530-2415.2009.01178.x

Prelec, D. (1998). The probability weighting function. *Econometrica*, *6*(3), 497–527.

Santos, L. R., & Rosati, A. G. (2015). The evolutionary roots of human decision making. *Annual Review of Psychology*, *66*, 321–347.

Sekścińska, K., Rudzinska-Wojciechowska, J., & Jaworska, D. (2021). Self-control and investment choices. *Journal of Behavioral Decision Making*, *35*(5), 691–705. https://doi.org/10.1002/bdm.2236

Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition*, *49*(1), 11–36.

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology*, *41*(1), 1–19. https://doi.org/10.1146/annurev.ps.41.020190.000245

Slovic, P. (1995). The construction of preference. *American Psychologist*, *50*(5), 364–371.

Somerville, J., & McGowan, F. (2016). Can chocolate cure blindness? Investigating the effect of preference strength and incentives on the incidence of Choice Blindness. *Journal of Behavioral and Experimental Economics*, *61*, 1–11.

Steenfeldt-Kristensen, C., & Thornton, I. M. (2013). Haptic choice blindness. *i-Perception*, *4*(3), 207–210.

Stewart, N., Chater, N., & Brown, G. D. (2006). Decision by sampling. *Cognitive Psychology*, *53*(1), 1–26.

Stewart, N., Chater, N., Stott, H. P., & Reimers, S. (2003). Prospect relativity: How choice options influence decision under risk. *Journal of Experimental Psychology: General*, *132*(1), 23–46.

Teal, J., Kusev, P., Heilman, R., Martin, R., Passanisi, A., & Pace, U. (2021). Problem gambling 'fuelled on the fly'. *International Journal of Environmental Research and Public Health*, *18*(16), 8607. https://doi.org/10.3390/ijerph18168607

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323.

Vlaev, I., Kusev, P., Stewart, N., Aldrovandi, S., & Chater, N. (2010). Domain effects and financial risk attitudes. *Risk Analysis*, *30*(9), 1374–1386.

von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior* (2nd ed.). Princeton University Press.

Wade, K. A., Green, S. L., & Nash, R. A. (2010). Can fabricated evidence induce false eyewitness testimony? *Applied Cognitive Psychology*, *24*(7), 899–908. https://doi.org/10.1002/acp.1607

Wang, J., Otgaar, H., Smeets, T., Howe, M. L., & Zhou, C. (2019). Manipulating memory associations changes decision-making preferences in a preconditioning task. *Consciousness and Cognition*, *69*, 103–112. https://doi.org/10.1016/j.concog.2019.01.016

Weiss-Cohen, L., Konstantinidis, E., & Harvey, N. (2021). Timing of descriptions shapes experience-based risky choice. *Journal of Behavioral Decision Making*, *34*(1), 66–84.

Zhang, Y., & Slovic, P. (2018). Much ado about nothing: The zero effect in life-saving decisions. *Journal of Behavioral Decision Making*, *32*(1), 30–37.

How to cite this article: Kusev, P., van Schaik, P., Teal, J., Martin, R., Hall, L., & Johansson, P. (2022). How false feedback influences decision-makers' risk preferences. *Journal of Behavioral Decision Making*, 1–9. https://doi.org/10.1002/bdm.2278