**BRIEF REPORT**

# Confidence in risky value-based choice

Kevin da Silva Castanheira [1] · Stephen M. Fleming [2] · A. Ross Otto [1]

## Abstract
Risk engenders a phenomenologically distinct experience from certainty, often driving people to behave in ostensibly irrational ways, and with potential consequences for our subjective sense of confidence in having made the best choice. While previous work on decision confidence has largely focused on ambiguous perceptual decisions or value-based choices under certainty, it is unclear how subjective confidence reports are formed during risky value-based choice (i.e. those with uncertain outcomes). Accordingly, we sought to examine the effect of risky (versus certain) choice upon confidence ratings in a calibrated economic choice task and explore the well-documented interrelationships between confidence and subjective value (SV) as well as choice response time (RT) in the context of value-based choice. By jointly analyzing choices (risky versus certain), SV of the chosen option, confidence, and RT, we found a systematic effect of risk on subjective confidence: subjective confidence reports were significantly higher when selecting a certain prospect compared with a risky one. Interestingly, risk attenuated the strength of the relationships between confidence and both RTs and difference in subjective value ($\Delta$SV), as well as the relationship between RT and $\Delta$SV. Taken together, these results corroborate how choice, RT, confidence and SV relate in value-based choice under risk, informing both theories of confidence and risk preferences.

In daily life, we often make risky choices about options with described risk levels, for which the resulting outcomes are uncertain. At the same time, our choices are accompanied by a subjective sense of confidence, which is meaningful, consequential, and thought to reflect a belief about having picked the best option in the absence of knowledge about the choice's outcome (Kiani, Corthell, & Shadlen, 2014). For example, behavioral economists find that investors' apparent overconfidence in their actions can, in part, explain the high failure rates of risky ventures (Camerer & Lovallo, 1999) as well as excessive trading in financial markets (Barber & Odean, 2001). In the laboratory, confidence judgments accompanying value-based choices (e.g., for food items) reflect an assessment of the decision-maker's accuracy in picking the subjectively superior option (De Martino, Fleming, Garrett, & Dolan, 2013), and accordingly, an individual reporting low

confidence is more likely to change their mind when faced with the same choice again (Folke, Jacobsen, Fleming, & Martino, 2016). Confidence can also guide future behavior due to its role in both controlling information seeking (Desender, Boldt, & Yeung, 2018), postdecisional processing (Rollwage et al., 2020), and learning by weighting incoming evidence (Meyniel & Dehaene, 2017; Meyniel et al., 2015). Thus, an accurate representation of confidence is in itself valuable, as it plays a role in monitoring errors and planning subsequent actions in the absence of feedback (Samaha, Switzky, & Postle, 2019).

As of late, there has been a surge of interest in studying decision confidence and the neural circuits underlying its computation in the domain of perceptual decision-making. Within a Bayesian framework, confidence has been characterized as the posterior probability of being correct given noisy sensory data (Aitchison, Bang, Bahrami, & Latham, 2015; Meyniel, Sigman, & Mainen, 2015). Interestingly, recent evidence suggests that individuals' confidence ratings, to varying extents, also incorporate a sense of certainty (or precision) about choice-relevant variables, over and above a perceived probability of being correct (Boldt, Blundell, & De Martino, 2019; Navajas, Bahrami, & Latham, 2016). Accordingly, a large body of work finds that confidence plays a useful role

✉ A. Ross Otto
  ross.otto@mcgill.ca

[1] Department of Psychology, McGill University, Montreal, QC, Canada

[2] Wellcome Centre for Human Neuroimaging and Department of Experimental Psychology, University College London, London, UK

in perceptual decisions in which uncertainty stems from sensory information.

However, in the context of value-based choice, examinations of subjective confidence reports have been constrained, up to now, to conditions of certainty: choosing between perceptually unambiguous concrete goods where the outcome is guaranteed (e.g., food items with measured subjective utility; De Martino et al., 2013; Folke et al., 2016; Polanía, Woodford, & Ruff, 2019). Surprisingly, little work has investigated how confidence is constructed in value-based choice under risk for which the resulting outcomes are probabilistic. Risky decisions engender a phenomenologically distinct experience (Loewenstein, Weber, Hsee, & Welch, 2001), which has long been thought to drive ostensibly irrational or inconsistent choice behavior (Tversky & Kahneman, 1992). For example, individual risk preferences are often inconsistent and context dependent: in choosing between a 'certain' prospect (guaranteed low payoff) and a 'risky' prospect (high payoff with low probability), an individual's risk preference depends on the magnitude of the outcome probability and the framing of the problem in terms of gains or losses (De Martino, Kumaran, Seymour & Dolan, 2006; Kahneman & Tversky, 1979). In such gambles, uncertainty does not stem from a lack of knowledge about the options themselves (i.e., ambiguity), as it would in a perceptual decision, but rather from the outcome of the choice. Importantly, risk and ambiguity have been previously shown to affect value-based choice behavior differently (Ellsberg, 1961; Tymula et al., 2012), but it remains unclear how both the presence of risk and the act of making a risky choice impact subjective confidence in the context of value-based decisions.

To address this question, we measured people's choices and confidence about risky prospects. Intuitively, we expected that choices made to certain prospects should be accompanied by an overall boost in subjective confidence, as these choices do not involve any risk with regards to the outcome. Further, we examine how risk alters the well-documented relationships between value, response time, and subjective confidence (De Martino et al., 2013; Folke et al., 2016). As decision-makers, under conditions of certainty, report higher confidence and faster RTs when the options have larger value differences (De Martino et al., 2013), we might expect that these relationships would be attenuated when a risky option is chosen,

given that risky choices are accompanied by an inherent uncertainty about the risky option's value. Similarly, as reported confidence tends to be higher for choices with faster RTs (Folke et al., 2016) we might expect that this relationship is also attenuated in risky choices, analogous to findings from perceptual decision-making in which the relationship between confidence and RT is tempered under greater sensory uncertainty (Kiani et al., 2014).

## Method

### Participants

Data were collected online using Amazon's Mechanical Turk; 300 U.S.-based volunteers (44% female) participated in the experiment for a base remuneration of $3.00 and a cash bonus, computed from the outcomes of two randomly selected trials. We collected a large participant sample to ensure (1) our analyses operated over a wide range of subjective value distributions, which were determined for each participant on the basis of their Calibration Phase data, and (2) we had adequate statistical power to detect potentially subtle and complex effects of risk upon the relationships between confidence-related variables.

It was critical that participants selected the risky option frequently enough in order to demonstrate that participants' choices were governed by the options' subjective values and not simply a default tendency to choose the certain option in accordance with their level of risk aversion. Accordingly, we first assessed participants' individual risk preferences in a choice calibration phase, which identified a number of risk-averse participants who would require inordinately large risky outcomes in order to offset their risk aversion and compel them to make risky choices for this task. Further, we used this calibration procedure to prescreen participants for which we could calibrate reasonable choice sets. Of the 300 initial participants, we screened out 109 risk-averse participants who, according to our modelling approach, required risky outcomes of $12 or larger (regardless of outcome probability) in order to select the risky option over the certain prospect of $1 (see Table 1). From these 191 eligible participants' test phase data, we further excluded participants if they met any of the following criteria: (1) chose correctly in fewer than 80% of the catch trials (7 participants); (2) inconsistent preferences as measured by a discrepancy in their parameter estimates between phases (3 participants); (3) failed to respond to more than 10% of trials (0 participants); (4) used less than 30% of the confidence scale (4 participants).

### Stimuli and procedure

All trials involved incentive-compatible decisions between gains (i.e. winning money) and were separated into two phases. In the first (calibration) phase, risky stimuli varied

**Table 1** Mean percentage risky choice for both phases and the average response times for each choice

| Phase | P(risky) (SE) | Choice | RT (SE) |
|---|---|---|---|
| Test | 0.358 (0.0107) | Certain | 1749 (5.65) |
| | | Risky | 1838 (7.63) |
| Calibration | 0.424 (0.0070) | Certain | 1666 (5.11) |
| | | Risky | 1750 (6.21) |

in the probability (either 0.5 or 0.9) of winning the non-zero outcome, each with five levels of differences in expected value—from −0.5 (favoring the certain option) to 0.50 (favoring the risky option), respectively. These 10 distinct stimuli were repeated 10 times for a total of 100 stimuli for the test phase (see Fig. 1a). After each response, we probed subjective confidence using a 7-point scale ("how confident are you that the choice you made was the best choice?") following previous work on value-based decision-making (De Martino et al., 2013).

For the second phase (test phase), differences in subjective-value (ΔSVs) were inferred based on choices made in the calibration phase, using a parametric prospect theory model (Tversky & Kahneman, 1992). We used this model to calculate the amount on offer for the risky option necessary to elicit a particular value of ΔSV between a certain prospect (which was always fixed to $1) and a risky prospect—with the same outcome probabilities as the test phase (see Fig. 1d). In other words, choices were between a certain prospect of $1 and a risky prospect with a probability (i.e., 0.5 or 0.9) of obtaining a variable amount of money, and zero otherwise (see Table S1 in the Supplementary Material for an example calibrated stimulus set). Following this calibration procedure, participants made choices in a test phase containing the calibrated choice sets, which was the focus of our analysis. For both phases of the experiment, the position of the prospects (left versus right side of display) were counterbalanced, and the order of presentation was pseudorandomized (each set of 10 stimuli was randomized before being presented one set at a time to avoid repeated trials).

## Subjective value estimation model

SV was calculated following a prospect-theory model (Equation 1), altered slightly to account for the idiosyncratic risk-seeking behavior of some participants observed during the calibration phase in pilot studies by adding a probability weighting parameter (see Equation 2), yielding an altered prospect theory (APT) model, as pilot studies revealed that participants' choices were better captured by this model than the standard prospect theory model (SPT). This pattern was echoed in our final sample as revealed by comparing the mean model likelihoods for the calibration phase (APT = 47.33, SPT = 63.69).

$$SV_{certain} = (V_{certain})^{\alpha} \tag{1}$$

$$SV_{risky} = p^{(1-\beta)} * (V_{risky})^{\alpha} \tag{2}$$

$$\Delta SV = SV_{risky} - SV_{certain} \tag{3}$$

$$p(Gamble) = \frac{1}{1 + e^{-\mu(\Delta SV)}}. \tag{4}$$

In this model, the probability of selecting one prospect over the other is dependent on the difference in subjective value.

Here, $p$ is the probability of obtaining the risky outcome and $V$ is the magnitude of the risky outcome. When the β parameter is 0, the probability remains unchanged; when it assumes the value of 1 (all information about probability is ignored); and as β increases from 0 to 1, probabilities are increasingly overweighed. This APT model (which includes the probability-weighing function) was used to estimate differences in subjective value for both the calibration and test phases. Parameter estimates for the model were inferred on an individual participant level using maximum likelihood estimation. We augmented the likelihood function with a penalty (or "pseudoprior") which discourages extreme parameter estimates and encodes our expectation on reasonable parameter values for these models, defined by Gaussian distributions imposed on the parameters μ, α ($M = 0$, $SD = 1$) and β ($M = 0.5$, $SD = 0.5$). Finally, exponential link functions were used to constrain the parameters μ and α to strictly positive values.

Note that choice sets in the test phase were determined individually based on the participant's parameter estimates for this model during the calibration phase. This way, we could elicit confidence ratings both risky and certain choices. See Table 2 for the median parameter estimates for the test phase. We also verified that the specified model was recoverable by conducting parameter recovery analysis on both the utility ($r = .894$), probability weighting function ($r = .814$) exponents, and the temperature parameter ($r = .84$; see Supplementary Material).

## Data analysis

We used a series of mixed-effects regressions to examine how choice-related variables altered confidence ratings and RTs. Specifically, three main models were estimated: (1) a model examining how ΔSV and risky versus certain choice (dummy coded as 1/0) jointly determine confidence ratings, (2) a model examining how choice RT and risky (versus certain) choice jointly predict confidence ratings, and (3) a model examining how ΔSV and risky (versus certain) choice jointly predict choice RTs. Further, to ensure confidence effects were not driven solely by the described risk level (i.e., the probability of winning for the risky choice) we also estimated a model jointly predicting confidence ratings from the categorical variables of outcome probability (either 0.5 or 0.9) and risky choice (i.e., risky or certain), in which both variables were effect-coded (−0.5 / 0.5). Finally, to control for any possible learning effects over the course of the experiment, all models included a linear predictor of trial number (see Tables 3, 4, and 5 for complete model coefficient estimates). RTs were log-transformed to remove skew, and all continuous predictor variables were Z-scored within participant and within risky choice. We excluded trials with RTs greater than three standard deviations with respect to each participant and choice type (i.e., certain versus risky). These models were estimated
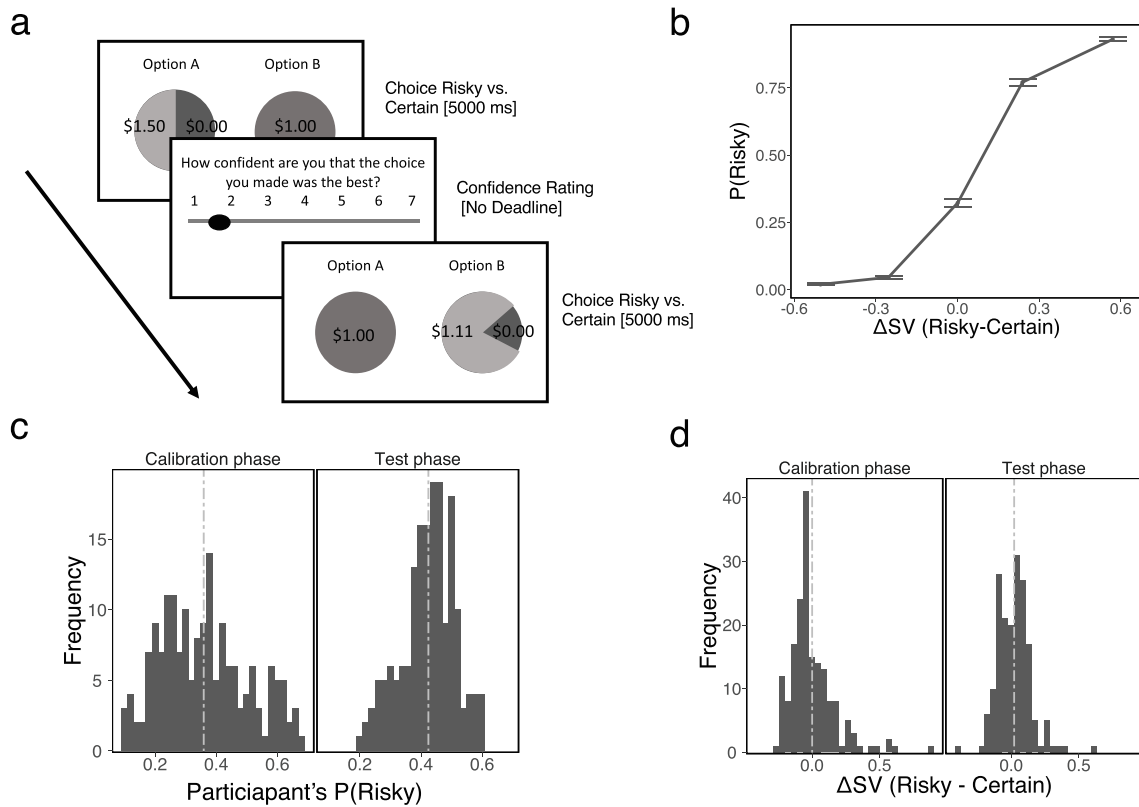
Fig. 1 Experimental Procedure. **a** In both the calibration and test phases, participants were asked to make a series of risky-value-based choices and rate their confidence on a 7-point scale. Using their choices in the calibration phase, stimuli were adjusted according to MLE parameter estimates such that their ΔSV distribution sampled a wide range and was centered at 0. **b** The probability of risky choice as a function of ΔSV with 95% confidence intervals for the test phase. Positive values of ΔSV indicate a preference for the risky option whereas negative values of ΔSV indicate a preference for the certain prospect. **c** The distribution of participants' overall proportions of risky choice per phase (test or calibration) with means plotted as dashed lines. Overall, the calibration was successful in tightening the distribution of participants' risk preferences. **d** The distribution of differences in subjective value for both phases (test or calibration) between risky and certain prospects, with means plotted as dashed lines

using the lmerTest package for R (Kuznetsova, Brockhoff, & Christensen, 2016), with all possible coefficients taken as random effects over participants, and employing Satterthwaite's method to compute *p* values.

individually calibrated each participant's choice set (specifically the money on offer for the given trial) in the main experiment. This ensured that the difference in subjective values between the risky and sure options (ΔSV) was centered around zero, effectively discouraging bias towards either risky or certain choices while also covering a range of values

## Results

Because a risk-seeking (or risk-averse) participant might overwhelmingly prefer the risky (or certain) option, we

**Table 2** Medians and standard deviations of the subjective-value model parameters used in the test phase

|       | Median | SD   |
|-------|--------|------|
| alpha | 0.65   | 0.33 |
| beta  | 0.30   | 0.24 |
| mu    | 6.77   | 3.01 |

**Table 3** Results of the linear mixed-effects model predicting confidence from Z-scored ΔSV

| Term | $\beta$ | SE B | t | p |
|------|---------|------|---|---|
| Intercept (certain) | 6.4295 | 0.0379 | 169.6376 | <.001 |
| SVZ (certain) | 0.3914 | 0.0192 | 20.4107 | <.001 |
| Risky choice | −1.0271 | 0.0721 | −14.2531 | <.001 |
| Trial | −0.0007 | 0.0002 | −3.2647 | .001 |
| SVZ × Risky Choice | −0.0573 | 0.0134 | −4.2670 | <.001 |

*Note.* Estimates of effect size are as follows: $r_m^2 = 0.23$ reflecting the contribution of the fixed-effect and $r_c^2 = 0.58$ reflecting the whole model

**Table 4** Results of the linear mixed-effects model predicting confidence from Z-scored log RT

| Term | $\beta$ | SE B | t | p |
|---|---|---|---|---|
| (Intercept) | 6.5285 | 0.0391 | 166.8730 | <.001 |
| RTZ | −0.2846 | 0.0295 | −9.6349 | <.001 |
| Risky choice | −1.0528 | 0.0720 | −14.6126 | <.001 |
| Trial | −0.0008 | 0.0002 | −3.5149 | <.001 |
| RTZ ×Risky Choice | 0.0630 | 0.0255 | 2.4691 | .014 |

*Note.* Estimates of effect size are as follows: $r_m^2 = 0.16$ reflecting the contribution of the fixed-effect and $r_c^2 = 0.52$ reflecting the whole model

favoring either option. This calibration procedure was successful as the stimuli presented reflected participants' underlying preferences and varied from favoring the certain prospect to favoring the risky prospect in accordance with the relative subjective values (ΔSV) of the two options (see Fig. 1b and Table 1).

These calibrated stimuli also meant that individual differences in risk preferences were largely minimized, resulting in a sample that was on average closer to being risk-neutral (see Fig. 1c). However, the calibration did not completely eliminate apparent risk aversion as evidenced by participants' asymmetrical choice function (see Fig. 1b; preference for the risky option as a function of ΔSV).

## Confidence and response times as a function of risky choice

As predicted, subjective confidence ratings accompanying certain choices were significantly higher than confidence ratings accompanying risky choices ($t = 54.014$, $p < .0001$; see Fig. 2a), which we term a certainty 'bonus'. Mirroring this pattern, risky choice RTs were significantly slower than certain choices ($t = 10.37$, $p < .0001$; see Fig. 2b).

**Table 5** Results of the linear mixed-effects model predicting log RT from Z-scored ΔSV

| Term | $\beta$ | SE B | t | p |
|---|---|---|---|---|
| Intercept (certain) | 7.4031 | 0.0151 | 489.8762 | <.001 |
| SVZ (certain) | −0.0379 | 0.0028 | −13.4553 | <.001 |
| Risky choice | 0.0441 | 0.0054 | 8.1931 | <.001 |
| Trial | −0.0004 | 0.0000 | −8.7585 | <.001 |
| SVZ × Risky Choice | 0.0111 | 0.0032 | 3.4750 | .001 |

*Note.* Estimates of effect size are as follows: $r_m^2 = 0.02$ reflecting the contribution of the fixed-effect and $r_c^2 = 0.50$ reflecting the whole model

## Confidence, subjective value, and response times

In line with previous research (De Martino et al., 2013; Folke et al., 2016), we found that the difference in subjective value between the two choices (ΔSV) predicted choice confidence such that larger differences in subjective values predicted higher choice confidence (β = 0.39, SE = 0.01, p < .001; see Fig. 2c and Table 3). Choice RTs also negatively predicted confidence, such that faster responses were accompanied by higher reported confidence (β = −0.28, SE = 0.02, p < .001; see Fig. 2d and Table 4). Similarly, larger ΔSV between the options predicted faster choice RTs (β = − 0.037, SE = 0.002, p < .001; see Fig. 2e and Table 5).

If risky choices were associated with a general reduction in confidence, we would predict that risky (versus certain) choice would exert a change in the intercepts, but not the slopes, of the relationships between ΔSV and confidence, between choice RT and confidence, and between ΔSV and choice RT. Intriguingly, however, the strength of the relationship between confidence and ΔSV, confidence and RT, and ΔSV and RT, were in fact attenuated under risky choice. After selecting the risky choice, we found ΔSV to be a worse predictor of confidence (Risk × ΔSV interaction, β = −0.057, SE = 0.013, p < .001; see Fig. 2c and Table 3). Similarly, while RTs were observed to (negatively) predict confidence, this relationship was attenuated when the risky option was chosen (Risk × RT interaction, β = 0.063, SE = 0.025, p = .014; see Fig. 2d and Table 4). Finally, we also found that risky choice significantly modulated the relationship between ΔSV and RT (Risk × ΔSV interaction, β = 0.011, SE = 0.003, p = .001; see Fig. 2e and Table 5). Below, we explore possible alternative explanations for the observed results.

While the calibration phase was designed to mitigate severe risk aversion, a subset of participants remained relatively risk averse. Thus, we sought to rule out the possibility that the observed effects of risky choice on confidence could be simply driven by these participants' risk aversion, such that the reduced confidence levels accompanying risky choices merely indicate that these choices go against their typical choices of the certain option. Accordingly, we identified a subset of participants (N = 58) as significantly risk averse based on their test phase choices (defined by making significantly less than 50% of risky choices; binomial test). We then reestimated the regression model predicting confidence as a function of risky choice and ΔSV without these participants. Critically, we found both the main effect of risky choice on confidence (β=-1.025, SE = 0.079, p < .0001) and the interaction between risky choice and ΔSV (β= − 0.0695, SE = 0.014, p < .001; see Table S2 in the Supplementary Material) held even after excluding these risk-averse participants, suggesting that our observed confidence effects did not merely stem from choices conflicting with a decision-maker's 'default' choice tendency. We further explored whether
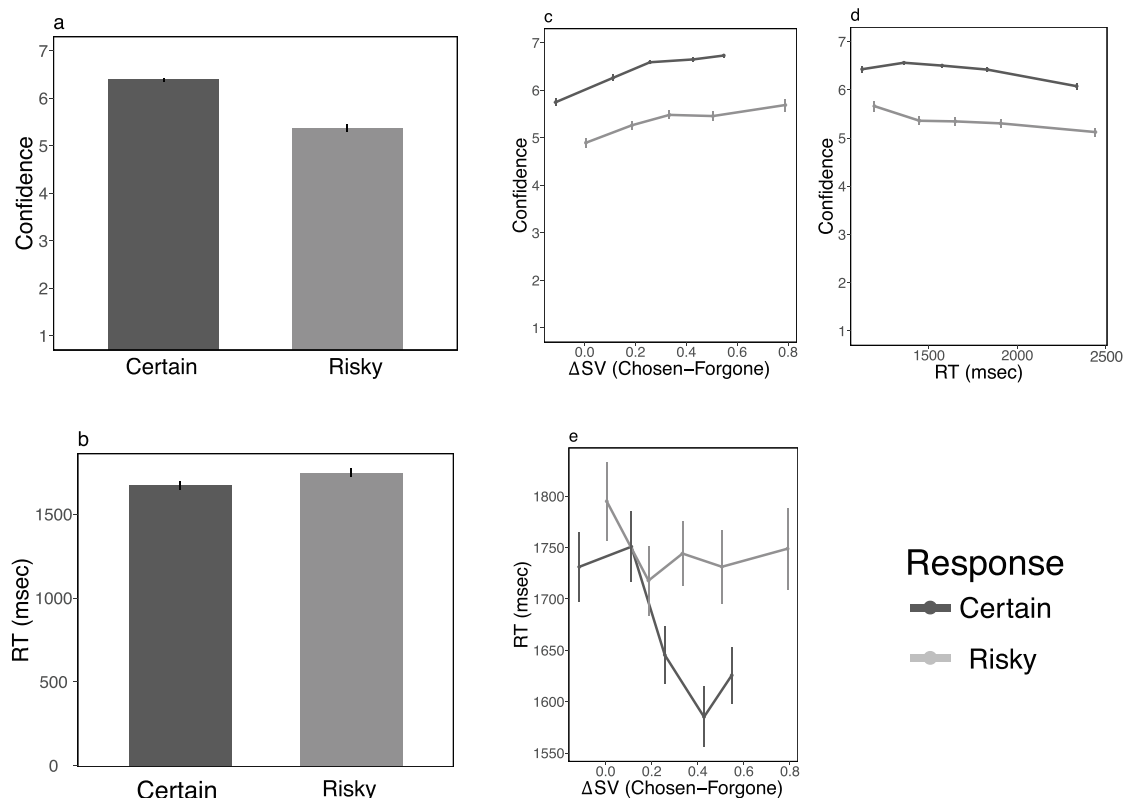
**Fig. 2** Effects of risky choice upon confidence ratings and response times (RTs). **a** Confidence plotted as a function of choice (risky versus certain). Certain choices engender higher overall confidence ratings compared with risky choices. **b** RTs plotted as a function of choice (risky versus certain). **c** Confidence as a function of ΔSV: Risky choice flattens the relationship between ΔSV and confidence when compared with certain choice. **d** Confidence as a function of RT: Risky choice also attenuates the relationship between RT and confidence. **e** RT as a function of ΔSV: Risky choice also attenuates the relationship between ΔSV and RT. Error bars represent 95% confidence intervals

individuals' underling risk preferences, indexed by the proportion of risky choices made in the calibration phase of the experiment, moderated the effect of risky choice on confidence. Accordingly, we estimated a mixed-effects regression predicting test phase confidence ratings from risky choice, proportion of risky choices made in the calibration phase, and their interaction. We failed to find a significant interaction ($\beta$ = 0.0268, $SE$ = 0.505, $p$ = .958; see Table S3 in the Supplementary Material), again suggesting that this certainty bonus did not merely reflect individual differences in risk preferences during the calibration phase.

Finally, given the observed effects of risky choice on confidence, we also tested whether these effects depended on the described risk level—the probability of winning. We ran a mixed-effects regression predicting confidence from outcome probability, risky choice, and their interaction (which were both effect coded) and observed a main effect of outcome probability ($\beta$ = 0.23, SE = 0.03, $p$ < .0001), such that subjects were more confident in picking the best option as $p$(win) increased. Additionally, we found a significant interaction between outcome probability (0.5 or 0.9) and risky choice ($\beta$ = 0.07, $SE$ = 0.03, $p$ = .01), indicating that the effect of described risk level upon confidence was stronger when making a risky

choice. Notably, the effect of taking a risky choice upon confidence was stronger than the effect of described risk level—for example, the mean difference in confidence between a risky and certain choice was approximately one point on the 7-point confidence scale (1.00), whereas the effect of outcome probability on confidence was considerably smaller (0.23). Finally, to confirm that the observed risk-induced decrease in confidence—and attenuation of the relationship between ΔSV and confidence—held when controlling for described risk level, we estimated a regression predicting confidence as a function of risky choice, outcome probability (dummy coded) and ΔSV, finding that both the observed decrease in confidence and attenuation of the relationship between ΔSV associated with risky choice remained significant (see Table S5 in the Supplementary Material).

## Discussion

What effect does the mere act of making a risky choice have on subjective confidence? While previous work examining subjective confidence reports about value-based choices has only considered choices made under certainty (De Martino

et al., 2013; Folke et al., 2016), here we directly compared the effects of risky versus certain choice on subjective confidence reports in value-based choice. We found that participants reported lower levels of confidence when they made risky compared with certain choices, which we interpret as reflecting decision-makers' internal state of uncertainty engendered by making a risky choice. Crucially, this confidence cost associated with risky choices was observed even in situations where the risky option had a larger subjective value (see Fig. 2c). In other words, the confidence levels associated with risky choices were lower than expected for a certain choice of equivalent subjective value.

Perhaps more notably, the mere act of making a risky choice appeared to attenuate interrelationships between ΔSV, RT, and confidence typically observed in value-based choice under certainty (De Martino et al., 2013; Folke et al., 2016). Specifically, we observed that the positive relationship between ΔSV and confidence, as well as the negative relationships between ΔSV and choice RT, as well as RT and confidence, were dampened when making risky choices. Our results complement this literature by suggesting that uncertainty, in the form of risk-taking, imparts fundamental changes to participants' subjective sense of confidence when making value-based decisions. Given that quality of evidence—operationalized here as the relative subjective value of the chosen option—as well as RT are theorized to play a role in the computation of subjective confidence (Kiani et al., 2014), it is conceivable that the intrinsic uncertainty of the risky option might fundamentally alter the computations relating RT and subjective value to reported confidence.

The results reported here also complement prior work examining subjective confidence in a range of tasks, not just value-based choice. This literature has debated whether confidence judgements reflect read-outs from the decision process at the time of choice (so-called decision locus theories) or depend on new information processed after the decision (postdecisional locus theories; Yeung & Summerfield, 2012), with mixed evidence for both accounts (Petrusic & Baranski, 2003). An open question raised by the present results is whether effects of risk on confidence occur at the time of choice (in accordance with the former view) or after the choice (in accordance with the latter). At first blush, the effect of risk on overall confidence levels (see Fig. 2a)—whereby the act of making a certain choice simply boosts the level of confidence reported after making a choice—appears compatible with both accounts, but the observed risk-evoked modulation of the relationship between RTs and confidence (see Fig. 2d) might be difficult to reconcile with a purely postdecisional boost. At any rate, the present results do not conclusively adjudicate between the possibilities of risk imparting a predecisional versus postdecisional bias in decision-making, which we highlight as important direction for future research.

Nonetheless this experiment provides an important initial demonstration of how risk-taking alters the computation of subjective confidence and highlights the usefulness of our choice set calibration procedure, which uniquely affords measurement of confidence—as a function of subjective value and RT—independent of individuals' risk preferences. Using this approach, we find that risk-taking disrupts the relationships between SV and confidence as well as RT and confidence typically observed in value-based choices (observed under certainty; e.g. De Martino et al., 2013). Understanding the computational origin of this risk-evoked bias could inform our understanding of both (1) how confidence is computed and used to guide future value-based choices, as previous work has found that confidence drives behaviors such as information seeking (Desender et al., 2018), confirmation bias (Rollwage et al., 2020), and evidence weighting during inference (Meyniel & Dehaene, 2017), and (2) risky decision-making more generally—for example, understanding how risk reduces decision confidence could help explain the ubiquitous nature of risk aversion (Kahneman, & Tversky, 1979).

It is worth noting that while our calibration procedure and exclusion criteria were motivated by model-fitting considerations (e.g., producing roughly even proportions of risky and certain choices), it is unclear whether the excluded subjects truly exhibited extreme risk aversion, or were influenced by other factors stemming from the choice sets used here. This potential constraint on the generalizability of our results should be addressed in future work, by looking across different participant populations and different payoff amounts. Alternatively, the observed risk preferences could have been specific to the reward incentives used (only the outcomes of two randomly chosen choices were realized) which raises the question of whether the observed effects of risk upon confidence could generalize to tasks in which every choice is nonhypothetical (Xu et al., 2018). Finally, it should be noted that we chose to calibrate and model choices using an altered prospect theory model, which assumes that outcome and probability information are integrated into a subjective value representation. Our choice of model was largely informed by the ubiquity of prospect theory, with recent work highlighting both its robustness (Ruggeri et al., 2020), and even its ability to successfully capture the use of simpler heuristics which do not integrate probability and outcome information. As the choice sets used here were designed under the assumptions of prospect theory, future work should aim to test the generalizability of our findings on confidence to choice sets amenable to the use of simpler choice heuristics, such as 'minimax' and 'maximax' (Pachur, Suter, & Hertwig, 2017).

# References

Aitchison, L., Bang, D., Bahrami, B., & Latham, P. E. (2015). Doubly Bayesian analysis of confidence in perceptual decision-making. *PLOS Computational Biology*, *11*(10), e1004519.

Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics*, *116*(1), 261–292.

Boldt, A., Blundell, C., & De Martino, B. (2019). Confidence modulates exploration and exploitation in value-based learning. *Neuroscience of Consciousness*, *2019*(1). https://doi.org/10.1093/nc/niz004

Camerer, C., & Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *The American Economic Review*, *89*(1), 306–318.

De Martino, B., Fleming, S. M., Garrett, N., & Dolan, R. J. (2013). Confidence in value-based choice. *Nature Neuroscience*, *16*(1), 105–110.

De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science*, *313*(5787), 684–687.

Desender, K., Boldt, A., & Yeung, N. (2018). Subjective confidence predicts information seeking in decision making. *Psychological Science*, *29*(5), 761–778.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics, 75*(4), 643–669. https://doi.org/10.2307/1884324

Folke, T., Jacobsen, C., Fleming, S. M., & Martino, B. D. (2016). Explicit representation of confidence informs future value-based decisions. *Nature Human Behaviour*, *1*(1). https://doi.org/10.1038/s41562-016-0002

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–291. https://doi.org/10.2307/1914185

Kiani, R., Corthell, L., & Shadlen, M. N. (2014). Choice certainty is informed by both evidence and decision time. *Neuron*, *84*(6), 1329–1342.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). lmerTest: Tests in linear mixed effects models [Computer software]. Retrieved from https://CRAN.R-project.org/package=lmerTest

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, *127*(2), 267–286.

Meyniel, F., & Dehaene, S. (2017). Brain networks for confidence weighting and hierarchical inference during probabilistic learning. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(19), E3859–E3868.

Meyniel, F., Schlunegger, D., & Dehaene, S. (2015). The sense of confidence during probabilistic learning: A normative account. *PLOS computational biology*, *11*(6). Public Library of Science.

Meyniel, F., Sigman, M., & Mainen, Z. F. (2015). Confidence as Bayesian probability: From neural origins to behavior. *Neuron*, *88*(1), 78–92.

Navajas, J., Bahrami, B., & Latham, P. E. (2016). Post-decisional accounts of biases in confidence. *Current Opinion in Behavioral Sciences*, *11*(Suppl. C), 55–60.

Pachur, T., Suter, R. S., & Hertwig, R. (2017). How the twain can meet: Prospect theory and models of heuristics in risky choice. *Cognitive Psychology*, *93*, 44–73.

Petrusic, W. M., & Baranski, J. V. (2003). Judging confidence influences decision processing in comparative judgments. *Psychonomic Bulletin & Review*, *10*(1), 177–183.

Polanía, R., Woodford, M., & Ruff, C. C. (2019). Efficient coding of subjective value. *Nature Neuroscience*, *22*(1), 134.

Rollwage, M., Loosen, A., Hauser, T. U., Moran, R., Dolan, R. J., & Fleming, S. M. (2020). Confidence drives a neural confirmation bias. *Nature Communications*, *11*(1), 1–11.

Ruggeri, K., Alí, S., Berge, M. L., Bertoldo, G., Bjørndal, L. D., Cortijos-Bernabeu, A., … Folke, T. (2020). Replicating patterns of prospect theory for decision under risk. *Nature Human Behaviour, 4,* 622–633. https://doi.org/10.1038/s41562-020-0886-x

Samaha, J., Switzky, M., & Postle, B. R. (2019). Confidence boosts serial dependence in orientation estimation. *Journal of Vision*, *19*(4), 25–25.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323. https://doi.org/10.1007/BF00122574

Tymula, A., Belmaker, L. A. R., Roy, A. K., Ruderman, L., Manson, K., Glimcher, P. W., & Levy, I. (2012). Adolescents' risk-taking behavior is driven by tolerance to ambiguity. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(42), 17135–17140.

Xu, S., Pan, Y., Qu, Z., Fang, Z., Yang, Z., Yang, F., … Rao, H. (2018). Differential effects of real versus hypothetical monetary reward magnitude on risk-taking behavior and brain activity. *Scientific Reports*, *8*(1), 1–9. https://doi.org/10.1038/s41598-018-21820-0

Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: Confidence and error monitoring. *Philosophical Transactions of the Royal Society, B*, *367*(1594), 1310–1321.