Epistemic and aleatory uncertainty in decisions from experience

Joel Holwerda (j.holwerda@unsw.edu.au)
School of Psychology
University of New South Wales
NSW, 2052, Australia

Ben R. Newell (ben.newell@unsw.edu.au)
School of Psychology
University of New South Wales
NSW, 2052, Australia

Abstract

People intuitively distinguish between uncertainty they believe is potentially resolvable and uncertainty that arises from inherently stochastic processes. The vast majority of experiments investigating decisions based on experience, however, have focused exclusively on scenarios that promote a stochastic interpretation by representing options as images that remain identical each time they are presented. In the current research, we contrasted this method with one in which the visual appearance of options was subtly differentiated each time participants encountered them. We found that introducing this variability to the appearance of options influenced the way people interpreted uncertainty. Although there was little evidence of an impact on exploration, these differences in interpretation may reveal other limitations to the generalisability of previous decision-making tasks.

Keywords: uncertainty; decisions from experience; exploration; ambiguity aversion; risky choice

Suppose that you were asked the question "what was John Lennon's middle name?"[1] and also asked to guess the outcome of a single roll of a six-sided die. Unless you happen to be an aficionado of Beatles' trivia, these questions would both entail some degree of uncertainty, but your perception of that uncertainty would likely differ considerably. Uncertainty about the trivia question would be interpreted as arising from a lack of knowledge whereas uncertainty regarding the die roll would be interpreted as resulting from an inherently stochastic process.

These interpretations reflect the two dimensions of uncertainty that Fox and Ülkümen (2011) suggest we intuitively differentiate. *Epistemic uncertainty* is attributed to insufficient information about a specific instance for which the truth is knowable, in principle, and is represented as confidence in one's knowledge or understanding of the underlying causal system. *Aleatory uncertainty*, in contrast, is perceived as arising from processes that are inherently stochastic, and as such, is represented as subjective probabilities over a class of possible outcomes. People generally use confidence statements ("sure", "certain") to refer to epistemic uncertainty and likelihood statements ("chance", "probability") to refer to aleatory uncertainty; in their use in natural language, these statements differ on a wide range of attributes including whether they refer to the past or future, the propensity to quantify uncertainty numerically, and the perceived level of control (Ülkümen, Fox, & Malle, 2016). Assuming that these linguistic differences either reflect or shape people's conceptualisation of uncertainty, this provides at least some evidence for a meaningful dichotomy.

The ability to determine whether outcomes are potentially knowable is particularly important when interacting with other agents—if a used car dealer offers to sell you a car at an unbelievable price, you might want to consider what they know and what you are missing. Avoiding options involving epistemic uncertainty might, therefore, offer a sensible strategy to avoid exploitation by an adversary with greater knowledge. Consistent with this, people strongly prefer gambles where probabilities are explicitly presented than ones where the odds are ambiguous (Ellsberg, 1963). They are also more likely to accept a gamble regarding stock prices in the future than the past (Heath & Tversky, 1991). Even children as young as four years old have been shown to be sensitive to whether outcomes have been resolved or whether they are yet to be decided (Robinson, Rowley, Beck, Carroll, & Apperly, 2006). Participants in each of these experiments seem to have made their decisions based on the amount they knew relative to the amount that they believed was knowable.

Responding to uncertainty is less straightforward, however, when choosing some options yields information that can reduce uncertainty when making future choices. When this is the case, epistemic interpretations—that uncertainty can be reduced by acquiring information—should lead to greater exploration than aleatory interpretations that uncertainty is irreducible. Furthermore, it makes sense to prefer options for which uncertainty is epistemic because acquiring information can increase future performance. Consistent with this, people are more likely to choose ambiguous options when they are presented repeatedly, which allows them to take advantage of epistemic uncertainty and improve their future choices (Liu & Colman, 2009). They are also more likely to search for patterns when the task is described as problem-solving as opposed to gambling (Goodnow, 1955). As such, in situations where it is possible to learn from experience, epistemic uncertainty is a double-edged sword; it indicates both the possibility that you might be less knowledgeable than others but also that your performance might improve over time.

So how do people solve this dilemma? The studies described above suggest that both aspects might influence behaviour but the impact of the interpretation of uncertainty has received very little attention regarding decisions made based

[1] For those playing at home, the answer is Winston.

on experience. This is likely because most prominent theories of decision-making assume an aleatory interpretation of uncertainty—they represent uncertainty as subjective probabilities over classes of events rather than specific instances (e.g., Kahneman & Tversky, 1979). As a result, the experiments designed to test these theories have almost exclusively used options that were represented by images that were identical each time they are presented (for a review, see Wulff, Mergenthaler-Canseco, & Hertwig, 2018). This is quite different from the decisions we face in our daily lives where variability in the outcomes is usually correlated with variability in the observed features of options.

There is an implicit assumption that decision-making experiments will generalise to these everyday situations but there is at least some evidence suggesting the contrary, that people interpret uncertainty as being more epistemic when they are presented with non-identical stimuli and that this, in turn, influences their choices (James & Koehler, 2011). As a concrete example, apples at a supermarket come in distinct varieties but each apple is subtly different from the others. The presence of this variability suggests that it might be possible to map the variability in the apples' sweetness (the uncertain outcome) onto their colour, shape, or size (the observed features), and therefore, the uncertainty regarding the sweetness of the apples is more likely to be interpreted as epistemic. If instead there was no variability in the observed features of apples—as is often the case with options in decisions from experience experiments—mapping the outcomes onto observed features would not seem possible, and therefore, uncertainty would be interpreted as aleatory.

In this paper, we manipulated the amount of variability in the observed features of options to examine whether people interpret uncertainty differently in standard decision-making tasks compared with situations that allow them to map uncertainty onto observed features of options. In doing so, we aimed to examine whether epistemic interpretations are more likely in situations with more naturalistic variability and determine whether people respond by avoiding or seeking epistemic uncertainty.

Experiment 1

The first experiment aimed to assess whether introducing more naturalistic variability to the appearance of options would influence the way people interpret uncertainty, as measured using the Epistemic–Aleatory Rating Scale (EARS) developed by Fox, Tannenbaum, and Ülkümen (2016). We expected that participants who made choices between pairs of safe and risky options represented by images that remained constant across trials, as is usually the case in a standard bandit task, would report higher aleatory and lower epistemic uncertainty associated with the risky option than participants who made choices between options that were subtly differentiated, more analogous to the choices made in daily life.

In both conditions, the safe option should be associated with low levels of uncertainty—both epistemic and aleatory—because the outcome only ever varies by a couple of points. As a result, disparity in the proportion of choices for the risky option between conditions may reflect the way participants are interpreting uncertainty regarding the risky option. We expect that participants will interpret uncertainty in a more epistemic fashion in the condition where options are differentiated, and therefore, if participants are averse to epistemic uncertainty, we would expect them to show a stronger aversion for the risky option in this condition, relative to the condition where options were identical. In contrast, if epistemic uncertainty leads to greater exploration, we would expect the opposite pattern, with participants preferring the risky option in the condition where each instance is slightly different because they are more likely to believe they can improve their future performance.

The results of Experiment 1a suggested the possibility that participants were responding to the EARS questionnaire with reference to the uncertainty in the appearance of the options rather than in the outcomes. Experiment 1b aimed to minimise that possibility by emphasising that the questionnaire referred to the outcome of a specific future choice. It also aimed to increase the likelihood that participants were correctly differentiating between the safe and risky options by paying them based on their performance and making the outcome distributions more distinct. These and other relevant differences between Experiment 1a and 1b are discussed below but the results are presented together because the observed patterns of behaviour were sufficiently similar.

Method

Participants A total of 240 undergraduate psychology students from UNSW Sydney participated in Experiment 1 (120 each in Experiment 1a and 1b). The average age was 19.3 years (SD = 2.4) and 175 participants were female. In addition to receiving course credit, participants in Experiment 1b were able to earn a small amount of money depending on their performance in the task (M = AU$5.67, SD = AU$1.48).

Design and procedure Participants completed the experiment in individual testing booths. At the beginning of the task, they were presented with written instructions that they would complete a task that required them to repeatedly make choices between pairs of options presented on a screen and that they should aim to earn as many points as possible. They were not given information about the distributions of points; instead they were required to learn about options by receiving feedback—the number of points received—about the selected option. Participants did not receive feedback on options that were not selected.

Participants completed 110 trials that each consisted of a choice between a low outcome-variance (safe) and a high outcome-variance (risky) option that had the same expected value. In Experiment 1a the outcomes of both options were drawn from a Gaussian distribution with a mean of 50 points. The standard deviation of the safe option was 1 point and the risky option was 20 points—this distribution was truncated

so that all outcomes were two-digit numbers between 10 and 90. The outcome distribution for the risky option in Experiment 1a was centred on the same mean as the safe condition, and therefore, 50 points was the most likely outcome for both options. To accentuate the different levels of risk associated with each of the options in Experiment 1b, a bimodal distribution that had peaks at 30 and 70 points was used for the risky option.

Participants were randomly allocated into two conditions that differed in the degree of variability in the appearance of each option across trials. In the constant image condition, the safe and risky options were differentiated by colour (red/blue) and by their position on the screen (left/right) and the images used to represent them remained identical across trials. In the unique image condition, the images were, likewise, differentiated by colour and position, but a slightly different image was used to represent each option one every trial (see Figure 1 for examples of these images). Importantly, the amount that the images varied was the same for the safe and risky options and although each stimulus was unique, the outcomes were still drawn from the same distribution as the options in the constant image condition, thus providing both conditions with the same amount of relevant information.

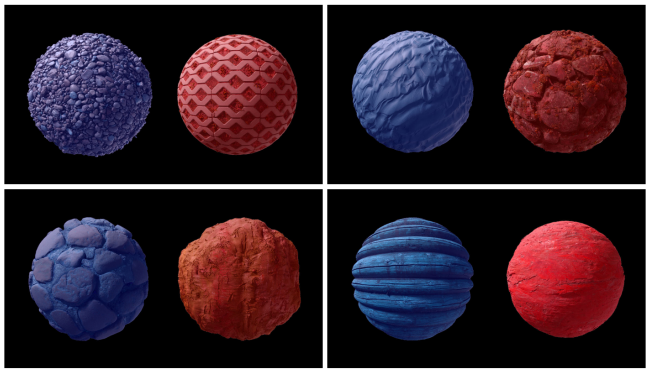

Figure 1: *Examples of pairs of images associated with the safe and risky options.*

Following the task, participants completed the ten-item EARS questionnaire with reference to the risky and safe options. In Experiment 1a, the participants were asked to think about the outcomes (numbers of points) that they received when they chose a red/blue option, responding to items on a seven-point Likert-scale that either indicated an epistemic interpretation (e.g., "The outcomes were in principle knowable in advance") or an aleatory interpretation (e.g., "The outcomes could play out in different ways on similar occasions").

The questionnaire in Experiment 1a was phrased in reference to past choices and referred collectively to the blue or red options—as opposed to a specific future choice—and it is possible that some participants responded with reference to the variability in the appearance of the options. Experiment 1b aimed to address this issue by presenting the epistemic uncertainty scale with reference to a future choice regarding a specific stimulus. Participants were asked to imagine that they were going to choose an option displayed on the screen and were presented with the epistemic uncertainty scale regarding the "outcome (number of points)" that would result from that choice.

Analysis All posterior distributions were determined by Hamiltonian Monte Carlo using the brms package in R (Bürkner, 2017). Weakly regularising priors were used for each parameter: for the logistic regression models, Gaussian distributions with a mean of 0 and standard deviation of 1 were used for the intercept and slope parameters and a half-Cauchy distribution with a location of 0 and scale of 1 was used for the standard deviation parameter in the hierarchical model of choices. For the uncertainty questionnaire responses, a Gaussian distribution with a mean of 0 and standard deviation of 3 was used for the intercept and slope parameters. A half-Cauchy distribution with a location of 0 and scale of 3 was used for the standard deviation parameters. All parameters had an effective sample size greater than 10000 and an $\hat{R} < 1.01$ indicating adequate chain convergence.

Results and discussion We hypothesised that participants who were presented with unique images on each trial would report higher epistemic and lower aleatory uncertainty associated with the risky option compared with participants who were always presented with the same images. To assess this, responses to the EARS questionnaire (see Figure 2) were analysed using Bayesian linear regression predicting mean responses to the aleatory and epistemic items, with experiment, condition (unique/constant), and uncertainty type (aleatory/epistemic) as fixed predictor variables and varying intercepts for each participant. Responses to the risky option were consistent with our hypothesis and the mean posterior estimate for the interaction between condition and uncertainty type was 0.74 (95% CI = 0.14, 1.33) suggesting that, compared with the standard bandit task, people interpret uncertainty as being more resolvable and less inherently stochastic when there is variability in the appearance of options.

Regarding the safe option, we expected that there would be little uncertainty and that responses to the uncertainty questionnaire would be similar across conditions. This was not what we observed. Responses to the EARS questionnaire for the safe option showed the opposite pattern as responses to the risky option and the mean posterior estimate for the interaction between condition and uncertainty type was -1.28 (95% CI = -1.90, -0.65). This suggests that participants interpreted the safe option as less epistemic and more aleatory when they were presented with a unique image on each trial.

One possible reason why this pattern emerged is that participants reported their uncertainty regarding the visual appearance of the options rather than their outcome. To address this, we designed Experiment 1b to emphasise that the EARS items refer to the outcome of a specific future choice, but the results were nearly identical to Experiment 1a, reducing credence in this explanation. Another explanation is

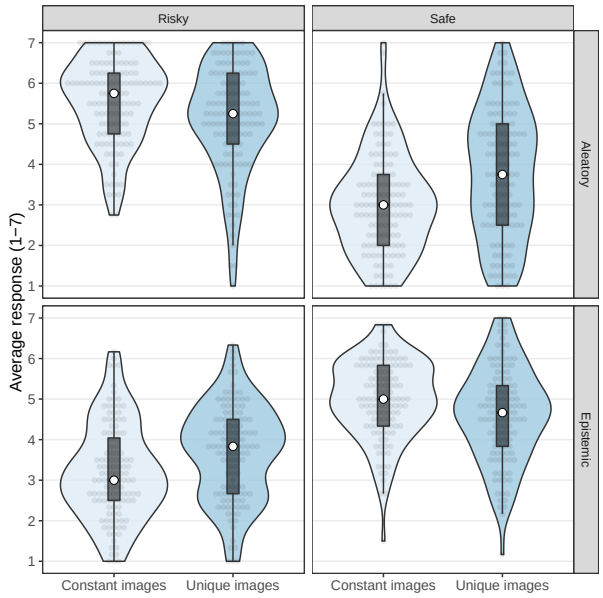

Figure 2: *Mean responses to the epistemic and aleatory subscales of the EARS questionnaire in Experiment 1.*

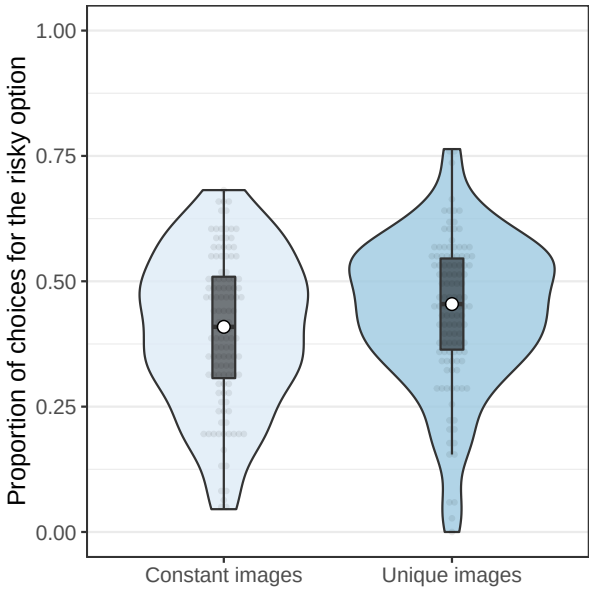

Figure 3: *The proportion of trials on which participants chose the risky option in Experiment 1.*

that variability in the appearance of options in our daily life is usually positively correlated with variability in their outcomes. Congruity between these two forms of variability may prompt epistemic interpretations that uncertainty is resolvable whereas a mismatch may increase aleatory interpretations that the outcomes are random.

The second main question we aimed to address in the first experiment was whether an epistemic interpretation would produce exploration or avoidance of the risky option. We examined this using Bayesian hierarchical logistic regression predicting the choice on each trial, with experiment and condition as fixed predictor variables and varying intercepts for each participant. Participants were six percent more likely to select the risky option when unique images were presented on each trial compared with when images remained constant (see Figure 3). This might indicate the possibility of a slight increase in exploration of the option with higher epistemic uncertainty, but the mean posterior log-odds for condition were 0.17 (95% CI = -0.09, 0.43) providing only weak evidence that this reflects an actual difference.

Experiment 2

The first experiment demonstrated that variability in the appearance of options influences the way uncertainty is interpreted but evidence regarding its influence on choices was far less definitive. There was some weak evidence that epistemic uncertainty leads to an increase in exploration but because participants were required to choose the option with higher epistemic uncertainty to gain the perceived benefits of exploration, it remains a possibility that aversion to epistemic uncertainty attenuated exploration in the task. The second experiment aimed to separate exploration from consequential choice, and therefore, obtain a clearer picture of whether epistemic uncertainty leads to exploration independent of potential aversion to epistemic uncertainty when making choices.

Method

Participants 137 undergraduate psychology students from UNSW Sydney participated in Experiment 2. The average age was 19.2 years (SD = 2.3) and 100 participants were female. In addition to receiving course credit, participants were able to earn a small amount of money depending on their performance in the task (M = AU$4.59, SD = AU$0.68).

Design and procedure Similarly to the first experiment, participants completed a decision task in which they were randomly allocated to either the constant image or unique image condition—the stimuli associated with each option and outcome distributions were identical to Experiment 1b. The key difference was that after choosing an option on each trial, instead of having the points both displayed on the screen and added to their total score, participants were required to decide whether they wanted to *observe* or *claim* the points associated with the option.

If they chose to observe, the number of points associated with the option was displayed on the screen but not added to their final score. If on the other hand, they chose to claim the outcome, the number of points was not displayed but was added to their final score and "points added to total" was presented on the screen. The distinction between observing and claiming was explained in detail prior to beginning the task and participants were told that they would complete a total of 100 trials. They were also required to obtain a perfect score on a short multiple choice questionnaire designed to ensure adequate knowledge of the task.

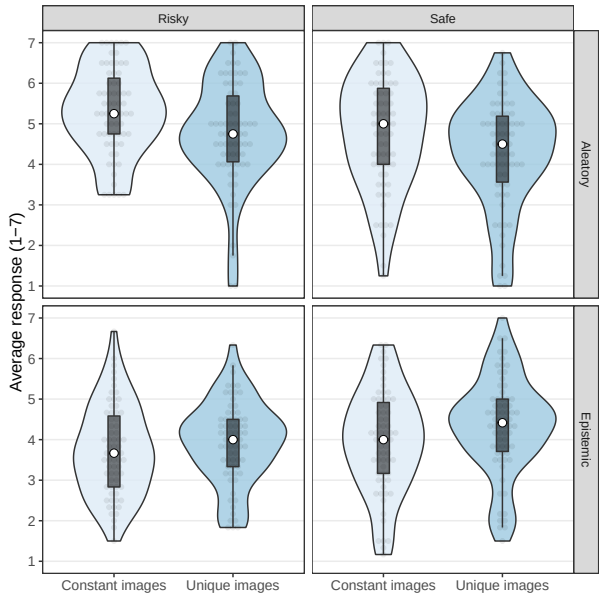

Figure 4: *Mean responses to the epistemic and aleatory subscales of the EARS questionnaire in Experiment 2.*

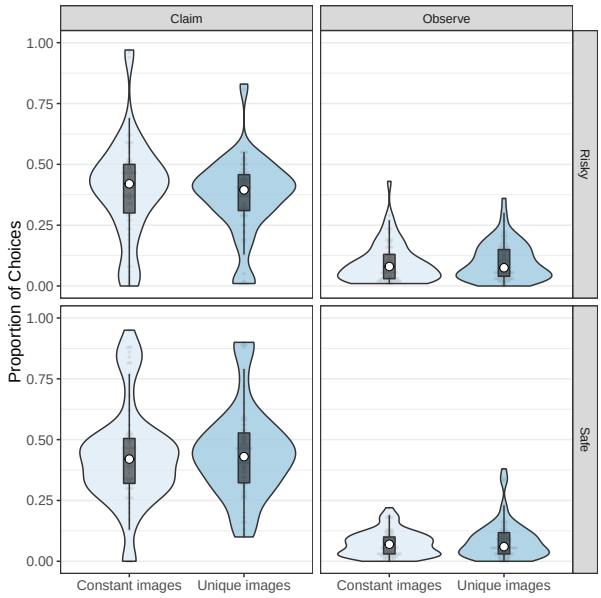

Figure 5: *The proportion of trials in Experiment 2 separated by whether the participant chose the safe or risky option and whether they chose to observe or claim the outcome.*

As was the case in the first experiment, participants were not given information about the outcome distributions and the only way to learn about options was to choose to observe the outcome. Doing so required them to forego the opportunity of claiming the points associated with the option, and therefore, it only made sense to observe an outcome if they believed it would provide information that could be exploited in future choices.

Analysis The same priors were used as in the first experiment. All parameters had an effective sample size greater than 10000 and an $\hat{R} < 1.01$ indicating adequate chain convergence.

Results and discussion Similarly to the first experiment, we were interested in whether variability in the appearance of options influences the interpretation of uncertainty. To address this question, we analysed responses to the EARS questionnaire (see Figure 4) using Bayesian linear regression predicting mean responses to the aleatory and epistemic items, with condition (unique/constant), and uncertainty type (aleatory/epistemic) as fixed predictor variables and varying intercepts for each participant. As we observed in the first experiment, participants who were presented with unique images on each trial reported higher epistemic uncertainty and lower aleatory uncertainty than those who were presented with the more standard bandit task: the mean posterior estimate for the interaction between condition and uncertainty type was 0.74 (95% CI = 0.20, 1.27).

In the first experiment, we unexpectedly found that the EARS questionnaire referring to the safe option produced the opposite pattern of responses as the risky option. We tentatively proposed that aleatory interpretations may arise when there is incongruity between the amount of variability in the appearance of the option and variability in the outcomes. We did not observe this pattern again in the second experiment and the mean posterior estimate for the interaction between condition and uncertainty type was 0.84 (95% CI = 0.31, 1.48). This pattern is quite similar to the one observed with the risky option and it is quite likely that participants were not differentiating between the riskiness of the options. Many participants explored each option less than five times, and therefore, they might not have perceived the mismatch in variability levels as unusual.

The second question we aimed to address was whether epistemic uncertainty would lead to higher exploration when doing so was no longer tied to consequential choice (see Figure 5). Because responses to the EARS questionnaire were similar for both options, we analysed overall exploration between the unique and constant image conditions. We used Bayesian hierarchical logistic regression predicting whether the participant claimed or observed the outcome, with condition (unique/constant) and choice (safe/risky) as fixed predictor variables and varying intercepts for each participant. The mean posterior log-odds for condition were -0.10 (95% CI = -0.46, 0.26) suggesting that there was a similar level of exploration regardless whether images were constant or unique on each trial. As mentioned above, many participants chose to observe the outcome of an option around five times during the task and chose to claim the outcome on the remaining trials. Although there was certainly no evidence for a large difference in exploration between conditions, it is possible that the low levels of exploration rendered the task insufficiently sensitive to detect differences across conditions.

General Discussion

The main purpose of the current study was to determine whether introducing subtle variability to the visual appearance of options would influence the interpretation of uncertainty in a task that required participants to learn from experience. The reason for doing so was that many of the decisions we make in our daily lives involve repeated interactions with options that are slightly different each time we encounter them whereas experiments studying decision-making exclusively present participants with options that are always identical. The experiments presented here demonstrate that people interpret uncertainty quite differently in these two contexts—variability precipitates a more epistemic interpretation of risky options—suggesting that we may need to investigate the degree to which findings from previous experiments can be generalised to everyday choices.

We also uncovered an unexpected finding that participants rated the safe option as higher in aleatory uncertainty when unique images were presented on each trial. One possible explanation is that the surface features of options are usually correlated with the underlying causal structure and variability in an option's appearance can be indicative of variability in its outcomes. This correspondence was violated in the first experiment and it might explain why uncertainty was perceived as more stochastic. It is also worth considering, however, that aleatory uncertainty was lower and epistemic uncertainty was higher for the safe option, which could indicate that aleatory items—such as "The outcome is unpredictable"—could also be capturing the total amount of uncertainty. It may be beneficial to integrate some measure of total uncertainty into future scales to investigate this possibility.

One seemingly plausible consequence of the differences in the interpretation of uncertainty is that it might influence the amount that people choose to explore options. If you think that your uncertainty could be resolved by acquiring more knowledge, it makes sense to explore. Despite this, neither experiment provided convincing evidence that increased epistemic uncertainty leads to increased exploration. Why might this be the case? Although there is a strong rational argument that epistemic uncertainty should lead to more exploration than aleatory uncertainty, there is surprisingly little experimental evidence supporting this claim. The relationship may be less straightforward than anticipated and people often use simple heuristics when performing exploration rather than conforming to rational solutions (Wilson, Geana, White, Ludvig, & Cohen, 2014); this might be the case with epistemic uncertainty.

Regardless, the current investigation into the interpretation of uncertainty in decisions based on experience has flagged an area that has been neglected because uncertainty has typically been treated as a unitary construct. The study of epistemic and aleatory interpretations of uncertainty is still in its infancy but there is evidence that people differentiate between them in several aspects of natural language (Ülkümen et al., 2016) and the consequences this distinction are slowly becoming apparent in areas such as the extremity of judgements (Tannenbaum, Fox, & Ülkümen, 2017). The absence of stark differences in the choices made between conditions in our research may provide reassurance about the generalisability of some findings that ignore epistemic uncertainty, but further inquiry may be required before its impact is fully realised.

References

Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1).

Ellsberg, D. (1963, 5). Risk, ambiguity, and the Savage axioms. In P. Gardenfors & N.-E. Sahlin (Eds.), *Decision, probability, and utility* (Vol. 77, pp. 245–269). Cambridge: Cambridge University Press.

Fox, C. R., Tannenbaum, D., & Ülkümen, G. (2016). The empirical case for distinguishing two dimensions of subjective uncertainty. *Working paper. University of California, Los Angeles, Los Angeles*.

Fox, C. R., & Ülkümen, G. (2011). Distinguishing two dimensions of uncertainty. *Perspectives on Thinking, Judging, and Decision Making*, 21–35.

Goodnow, J. J. (1955, 3). Determinants of Choice-Distribution in Two-Choice Situations. *The American Journal of Psychology*, *68*(1), 106.

Heath, C., & Tversky, A. (1991, 1). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty*, *4*(1), 5–28.

James, G., & Koehler, D. J. (2011). Banking on a bad bet: Probability matching in risky choice is linked to expectation generation. *Psychological Science*, *22*(6), 707–711.

Kahneman, D., & Tversky, A. (1979, 3). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263.

Liu, H.-h., & Colman, A. M. (2009). Ambiguity aversion in the long run : Repeated decisions under risk and uncertainty. *Journal of Economic Psychology*, *30*(3), 277–284.

Robinson, E. J., Rowley, M. G., Beck, S. R., Carroll, D. J., & Apperly, I. A. (2006, 11). Children's Sensitivity to Their Own Relative Ignorance: Handling of Possibilities Under Epistemic and Physical Uncertainty. *Child Development*, *77*(6), 1642–1655.

Tannenbaum, D., Fox, C. R., & Ülkümen, G. (2017). Judgment Extremity and Accuracy Under Epistemic vs. Aleatory Uncertainty. *Management Science*, *63*(2), 497–518.

Ülkümen, G., Fox, C. R., & Malle, B. F. (2016). Two dimensions of subjective uncertainty: Clues from natural language. *Journal of Experimental Psychology: General*, *145*(10), 1280–1297.

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore–exploit dilemma.

Journal of Experimental Psychology: General, *143*(6), 2074–2081.

Wulff, D. U., Mergenthaler-Canseco, M., & Hertwig, R. (2018). A meta-analytic review of two modes of learning and the description-experience gap. *Psychological Bulletin*, *144*(2), 140–176.