Title
Invaraince of Information Seeking Across Reward Magnitudes

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

ISSN
1069-7977

Authors
Liew, Shi Xian
Newell, Ben

Publication Date
2021

Peer reviewed

Invariance of Information Seeking Across Reward Magnitudes

Shi Xian Liew (s.liew@unsw.edu.au)

Ben R. Newell (ben.newell@unsw.edu.au)
School of Psychology, UNSW
Sydney, NSW 2052, Australia

Abstract

Most theoretical accounts of non-instrumental information seeking suggest that the magnitude of rewards has a direct influence on the attractiveness of the information. Specifically, the magnitude of rewards is assumed to be proportional to the strength of information seeking (or avoidant) behaviour. In a series of experiments using numerical and pictorial stimuli, we explore the extent to which observed information seeking behaviour tracks these predictions. Our findings indicate a robust independence of information seeking from outcome magnitude and valence with preferences for information largely remaining constant across different reward valence and magnitudes. We discuss these results in the context of current computational models with suggestions for future theoretical and empirical work.

Keywords: information seeking; reward magnitude; reward valence;

Introduction

Information is often sought out to guide action under risky situations. For instance, the avid poker player may look for specific patterns of behaviour in their opponent (i.e., a tell) to gauge whether the opponent may be bluffing. However, information about uncertain outcomes can also be attractive even when it appears non-instrumental in the decision-making process, that is, when the information cannot be used to guide any decision relevant to the task. For instance, a voter may be keen to follow the news on how the candidates are performing even though such news cannot change their vote (because it was already cast). Roulette players continue to eye the moving ball and wheel even though such information (i.e., ball trajectory and wheel speed) cannot change the bets already placed for that game. If information is valued only for its instrumentality to making task-relevant decisions, then in neither of these cases should the decision-maker expend resources to seek out such information.

Empirical evidence demonstrating non-instrumental information seeking behaviour has been robustly produced in a number of different studies (Sharot & Sunstein, 2020; Vasconcelos, Monteiro, & Kacelnik, 2015). The strength of such behaviour appears to depend on a number of key features of risky choice, including the probability of rewards (Iigaya et al., 2020; Charpentier, Bromberg-Martin, & Sharot, 2018), the delay between the presentation of an informative cue and the outcome (Iigaya, Story, Kurth-Nelson, Dolan, & Dayan, 2016; Embrey, Liew, Navarro, & Newell, 2020), and the valence of the outcome itself (i.e., whether the outcome is desirable or repulsive; Charpentier et al., 2018; Zhu, Xiang, & Ludvig, 2017).

However, not all features of risky choice tasks have been investigated. Notably, the effect of outcome *magnitude* on information preference has received minimal attention. Perhaps the closest empirical investigation into possible effects of outcome magnitude is seen in Bennett, Bode, Brydevall, Warren, and Murawski (2016), who used a task where participants gradually revealed (sometimes costly) cues on uncertain outcomes. Bennett et al. (2016) found that increasing the cost of informative cues directly reduced the preferences for such advance information. The cost of informative cues was not explicitly presented, instead they were shown as reductions in the eventual reward on each trial (e.g., obtaining \$2 on a winning trial instead of \$3 due to the the information cost of \$1). These results may hint towards a positive relationship between the eventual outcome magnitude and information seeking behaviour.

The lack of empirical work investigating the relationship between outcome magnitude and information seeking is striking for a number of reasons. First, outcome magnitude is a core feature of risky events, without which it is impossible to determine the expected value of the event. Assuming that people are cognizant of outcome magnitudes when engaging in risky choices, it does not seem a stretch to assume that these magnitudes play a role in how information on those risky choices are perceived. Second, current theories of information seeking invariantly include outcome or reward magnitude in their computation of information preferences. According to the framework described by Sharot and Sunstein (2020), people seek out non-instrumental information for its hedonic (or affective) utility. Within this framework, more positive outcomes (i.e., rewards of greater magnitude) can result in greater hedonic utility, and consequently greater seeking of information. More formally, the Reward Prediction Error with Anticipation model (RPE-A; Iigaya et al., 2016) assumes that information seeking is a function of the anticipation of positive outcomes—consequently, more positive outcomes (higher reward magnitude) would result in greater anticipation and thus more information seeking.

The present work aims to provide a rigorous exploration of how different outcome magnitude structures can affect information seeking behaviour. We first discuss three empirically-tested models of human information seeking (including RPE-A) and their qualitatively distinct patterns of outcome-dependent information seeking in more detail in the next section, while also considering two additional models as a theoretical baseline. Following this, we present two novel experiments investigating magnitude-dependency in humans.

Information Seeking Models

Instead of providing the technical specifications of each model in their entirety, we briefly indicate the central assumptions of each model and their relevant treatment of reward outcomes. For each model we identify the core reward-integration function that is proportional to the actual predictions of the model—these functions serve as concise summaries of model behaviour.

The Reward Prediction Error-Anticipation (RPE-A) model designed by Iigaya et al. (2016) is a reinforcement learning model that assumes information preferences result from the savouring of information about positively valenced outcomes, and conversely, information avoidance results from the dread experienced from information about negatively valenced outcomes. If a gamble has two strongly positive outcomes (e.g., gaining either \$100 or \$500), RPE-A would predict higher information preferences than a gamble with two weakly positive outcomes (e.g., gaining either \$1 or \$5). The predictions of RPE-A (i.e., the probability of choosing to seek information) are essentially proportional to the average reward magnitude from both outcomes:

$$Pr(\text{Info}) \propto \frac{r_w + r_l}{2}$$

where r_w and r_l indicate the rewards from a winning and losing outcome respectively. Since the denominator is a scaling constant, we can simplify the expression:

$$Pr(\text{Info}) \propto r_w + r_l \tag{1}$$

For our purposes here, we assume that every gamble has equal outcome probability (i.e., probability of winning and losing is .5).[1]

The Anticipated Prediction Error model (APE; Zhu et al., 2017) is also an anticipation-based reinforcement learning model, but instead of being driven by the savouring of anticipated futures, it assumes information seeking is the result of anticipatory signals from different attentional weights to the winning and losing outcomes. Crucially, APE only considers the absolute quantity of the mean reward value. It predicts the same amount of information preference for a risky choice between gaining \$100 or losing \$100 and gaining \$200 or losing \$200. More formally, we can define its information preference as proportional to the absolute value of the mean (or equivalently, the total) outcome value:

$$Pr(\text{Info}) \propto |r_w + r_l| \tag{2}$$

[1]This suggests that Equation 1 can be more generally expressed as the expected value of the gamble. However, since varying probabilities is beyond the scope of the present study, and to maintain a level of comparability of this equation across all models, we keep this special formulation here.

Bennett et al. (2016) formulated the Uncertainty Penalty model (UP) which assumes that people seek information to resolve the uncertainty inherent in risky choices. Unlike RPE-A, UP is agnostic to the valence of the rewards themselves and consequently does not predict information avoidant behaviour. Unlike APE which takes the absolute quantity after the summation of outcomes, UP's predictions are proportional to the summation of the absolute magnitudes of outcomes:

$$Pr(\text{Info}) \propto |r_w| + |r_l| \tag{3}$$

Model Predictions

To allow for a comprehensive comparison of different magnitude structures, we define a series of gamble conditions each with 50% probability of obtaining the winning or losing outcome. We vary the outcomes such that they can take either one of two magnitudes a or b where $a < b$, and can be either positive or negative in valence (i.e., a or $-a$). No condition can include outcomes with both equal magnitudes and valences. This results in six possible conditions, described as pairs of outcomes with the winning outcome followed by the losing outcome: 1) $-a, -b$; 2) $a, -a$; 3) $a, -b$; 4) $b, -a$; 5) $b, -b$; and 6) b, a.

Across the six gamble conditions, each model produces a distinct pattern of predictions. For instance, while RPE-A would predict lowest information seeking in condition 1 (where $r_w + r_l = -a - b$) and highest in condition 6 ($r_w + r_l = a + b$), UP would predict lowest information seeking in condition two ($|r_w| + |r_l| = a + |-a| = 2a$) and highest in condition 5 ($|r_w| + |r_l| = b + b = 2b$). We present these diagnostic predictions for all conditions in Figure 1. Note that RPE-A is unique in predicting information avoidant behaviour when all outcomes are negative (i.e., condition 1). While these predictions were directly generated by the models, we present them here at a qualitative rather than a quantitative level (e.g., RPE-A's probability of seeking information at condition 6 would be highest relative to the other conditions, but it need not be strictly close to 1.0).

For theoretical rigour we also consider other simple models that adopt orthogonal assumptions to the current models. All current models assume that some combination of magnitude has a proportional impact on information seeking, and among these models only RPE-A assumes that outcome valence can change the direction of information seeking. We can consider a model that does the opposite, that is, it ignores the absolute values of magnitude while allowing the polarity (or sign) of valence to drive the direction of information seeking. Like RPE-A, this sign-only model would predict lowest information preference at the condition 1, and highest preference at condition 6, but with every other condition producing intermediate information preference (Figure 1, second column from the right). More formally, we can express this sign-only model's predictions as a proportion of the average valences of the outcomes:

$$Pr(\text{Info}) \propto \frac{r_w}{|r_w|} + \frac{r_l}{|r_l|} \tag{4}$$

Table 1: Core Assumptions of Information Seeking Models on Outcome Values

Model	Valence	Magnitude	Mechanism
RPE-A	Yes	Yes	$r_w + r_l$
APE	No	Yes	$\lvert r_w + r_l \rvert$
UP	No	Yes	$\lvert r_w \rvert + \lvert r_l \rvert$
Sign-Only	Yes	No	$\frac{r_w}{\lvert r_w \rvert} + \frac{r_l}{\lvert r_l \rvert}$
Null	No	No	-

For completeness, we also consider a null model where neither outcome valence nor magnitude have any effect on information seeking. We summarise all core model assumptions and features in Table 1.

Experiment 1

To observe how information seeking behaviour changes with different outcome magnitudes, we adopt the *secrets* task also used in Iigaya et al. (2016) and Embrey et al. (2020). In this task, participants on each trial are given the option to either "Find Out Now" (FON) about the outcome of a gamble, or "Keep it Secret" (KIS), after which they receive the outcome (a gain or loss of points) after a fixed delay. Outcome magnitudes were manipulated to follow the six-condition structure described for model predictions.

Method

Participants We recruited 49 people (21 females, 28 males, $M_{age} = 34.02$ years) via the Amazon Mechanical Turk platform. Participants were compensated with 6.00 USD for participating in the task and could earn bonus amounts depending on the gamble outcomes on each trial ($M_{bonus} = 3.00$ USD). Bonus amounts were converted from the total points accumulated at the end of the session at the rate of 1.00 USD per 1000 points.

Design and Procedure On each trial, participants were presented with a gamble between two equally probable outcomes (of different magnitudes) and could choose to either receive information about the gamble outcome immediately or after a 20 second delay, with the clarification that their choices did not affect the outcome of the gamble. Choosing to FON presented participants with either a smiling face in the event of a winning outcome or a frowning face otherwise. Choosing to KIS presented participants with a confused-looking face. No matter their choice, participants had to wait 20 seconds after cue presentation before receiving the reward. The total reward value accumulated thus far was visible to participants on every trial. Participants faced six conditions, with ten trials in each condition. The conditions differed from each other only in outcome values according to the structure defined earlier for model predictions, with a fixed to 100 points and b set to 500 points (e.g., in condition 1 participants could win either 100 points or lose 500 points). The overall task procedure is presented in Figure 2. Participants were provided with a starting score of 3000 points to minimise the risk of finishing the task with a negative score.

Results and Discussion

Overall, a Bayesian t-test indicated decisive evidence[2] for information seeking ($M = .59, BF_{10} = 1.52 \times 10^{17}$; Figure 3, upper panel). However, a Bayesian ANOVA indicated decisive evidence for no differences in information preference between conditions ($BF_{01} = 936.64$). A visual comparison of these results in the upper panel of Figure 3 with the qualitative predictions in upper panels of Figure 1 suggests that the null model offers the closest approximation.

It is possible that the observed invariances in information preferences is due to the perceptual similarities in the magnitudes across conditions. That is, participants may not have viewed quantities such as "100" as being entirely different to "-100" or "500" (at least, for the purposes of evaluating information about the gamble).

Experiment 2

In Experiment 2, we use pictorial depictions of the outcomes to increase the perceptual salience of the rewards, hypothesising that this increase in outcome salience can lead to stronger differences in processing reward magnitudes, resulting in corresponding differences in information seeking behaviour. We also increased the number of trials per condition while reducing the number of conditions to allow for a more powerful test for any possible differences between conditions.

Method

Participants Participants were 50 people (18 females, 32 males, $M_{age} = 39.34$ years), screened to ensure they had not participated in Experiment 1. Compensation procedures were identical to those in Experiment 1.

Design and Procedure Experiment 2 used a near-identical design and procedure to Experiment 1 with three notable exceptions. First, to increase the salience of the reward magnitudes, the outcomes were presented as graphics instead of numbers. Each gain of 100 points was indicated by a picture of one moneybag, and each loss of 100 points was indicated by a cartoon picture of one robber (see Figure 2). Second, we use three conditions in Experiment 2 (instead of six previously). We include conditions 1 and 6 from Experiment 1, and add a new condition with a separate quantity $c, -c$ where c is fixed to 200 points. To clarify, the three conditions in Experiment 2 are 1) $-a, -b$; 2) $c, -c$; 3) b, a. Third, 15 trials of each condition were run (instead of 10).

[2]Our interpretations of Bayes factors follow the guidelines set forth by Kass and Raftery (1995)

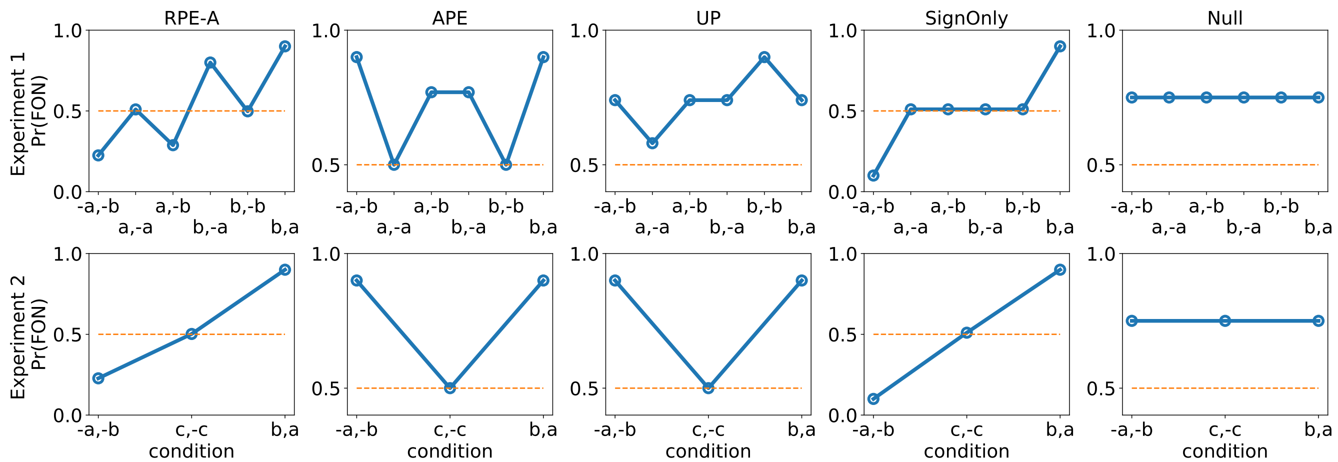

Figure 1: Qualitative predictions of information seeking models across varying outcome magnitude conditions. Blue markers and lines indicate model predictions, orange discontinuous line indicates Pr(FON) = .5.

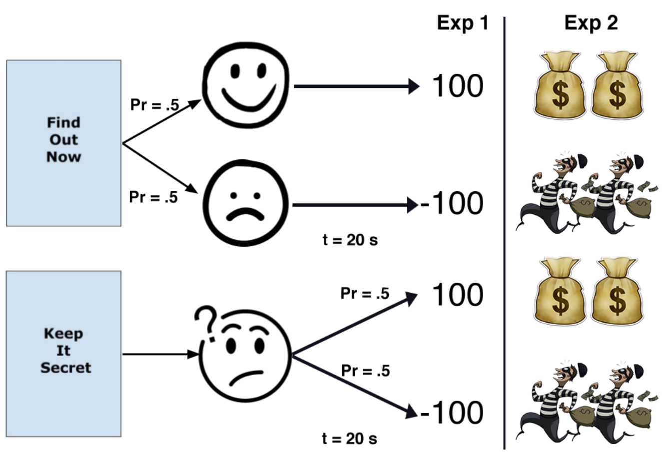

Figure 2: General experimental design. On each trial participants selected either Find Out Now or Keep It Secret. The former option revealed a cue indicating the outcome, while the latter option revealed an ambiguous cue. The outcome is presented 20 seconds after the cue. Experiment 1 outcomes were presented in numerical points, Experiment 2 outcomes were presented using graphical representations of points (1 moneybag = 100 points, 1 robber = -100 points).

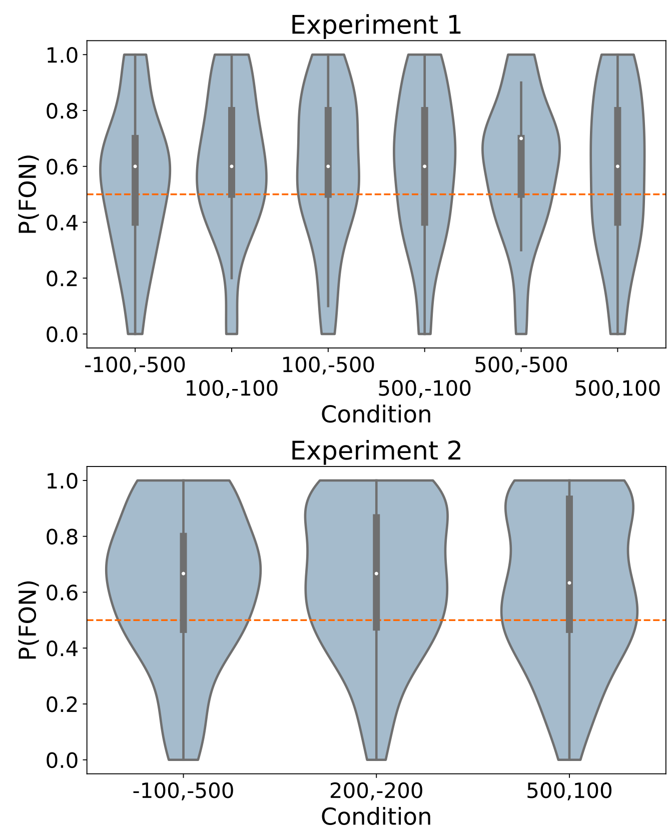

Figure 3: Violin plots of information seeking behaviour for Experiment 1 (upper panel and Experiment 2 (lower panel). Orange discontinuous line indicates the level where Pr(FON) = .5

Results and Discussion

A Bayesian *t*-test indicated decisive evidence for information seeking on average ($M = .64, BF_{10} => 10^{20}$; Figure 3, lower panel). A Bayesian ANOVA again indicated strong evidence for no differences in information preference between conditions ($BF_{01} = 96.27$).

Despite increasing the perceptual salience of the stimuli and increasing the number of trials, information seeking behaviour did not seem to change across across gambles with different outcome magnitude structures. Similar to Experiment 1, a visual comparison of these results in Figure 3 (lower panel) with the corresponding qualitative predictions in the lower panels of Figure 1 suggests that the null model offers the closest approximation of the data.

General Discussion

Our experiments found that varying the magnitude of gamble outcomes did not meaningfully alter information seeking behaviour, challenging the various assumptions of current models. Qualitatively, the best account we have for our data appears to be the null model, which suggests that neither reward valence nor magnitude factor into people's decision processes about information preference.

The observed invariance in information seeking across different magnitude structures appears to be the norm rather than the exception. We observed this in our experiments using both numerical as well as graphical outcomes. Further, the results of Experiment 2 closely resemble the findings reported by Embrey et al. (2020) who ran a comparable study using primary reinforcers. Primary reinforcers can be understood as rewards that can be immediately enjoyed and contrasts with secondary reinforcers, which are rewards provided as a proxy for primary reinforcers (e.g., money, a secondary reinforcer, being used as a medium to buy food, a primary reinforcer). More specifically, Embrey et al. (2020) exposed participants to a similar information seeking task as the present study, but offered chocolates as a positively valenced stimulus and an aversive microphone feedback sound as a negatively valenced stimulus. Across three conditions analogous to Experiment 2 of the present study (i.e., one where participants either won a chocolate or nothing, one where they either won a chocolate or received the aversive sound, and one where they either received the sound or nothing), Embrey et al. (2020) found consistent levels of information seeking averaging around 60% of choices, similar to the values found in the present study.

Our results initially appear to contradict the findings of Bennett et al. (2016), who found that increasing the cost of information (thereby decreasing the eventual reward) made people less likely to seek out information. However, we contend that the present study is investigating a fundamentally different effect. To clarify, in the task by Bennett et al. (2016), the actions taken by the decision-maker (to receive advance information or not) directly affected the expected reward in the task, creating differences in outcome magnitude *within* a trial. Decision-makers were always able to avoid the cost of information if they so wished. Consequently, the decrease in information seeking preferences can be the result of maximising the objective expected reward (i.e., choosing not to know because it has highest reward) as opposed to receiving less hedonic utility from a lower-valued outcome. In contrast, in the present study the choices made by the participant had no effect on the expected reward, and only differences in outcome magnitude *between* trials were analysed. Participants could not avoid changes in outcome magnitude from one trial to another, so any changes in their information preferences must be due to how they perceived outcome magnitude and not the result of maximising objective rewards.

Ignoring magnitude may appear to be an implausible assumption—not only do all of the current models explicitly consider magnitude (albeit with different mechanisms), but it may seem counter-intuitive to suggest that people care equally about advance information when considering small versus large rewards. Before addressing *why* this is occurring it may be worth considering *how* it is occurring. One possible mechanism for this behaviour may be that decision-makers perform feature normalisation when considering information from risky choices. Seen more commonly in the machine learning literature (e.g., Ekenel & Stiefelhagen, 2006; Aksoy & Haralick, 2001) feature normalisation is a process of transforming stimuli values onto a common scale. For instance, min-max normalisation takes the set of stimuli values and proportionally scales it such that the minimum and maximum values are at some predefined boundaries. What is preserved after this process is the relative differences in stimuli values and not the absolute values they once contained.

The specific reason as to why people seem to ignore magnitudes (whether by feature normalising or not) still remains unclear. It is possible that the outcome magnitudes used here were simply not different enough from one condition to another. Although rewards were presented in the order of hundreds of points, participants may have been converting the points to their actual monetary value and consequently not see much difference between winning/losing \$0.10 on one trial and winning/losing \$0.50 on another. Perhaps differences in information preference can be expected to arise when considering gambles that are orders of magnitude apart; for example, having one gamble with equally probable outcomes of \$1 and \$2 versus another with outcomes of \$1000 and \$2000.

On another level, future work could consider systematically exploring how information seeking is affected by *within*-gamble outcomes (and consequently, its certainty). For instance, it is reasonable to expect that people would be much less motivated to seek advance information on gambles which have very similar outcomes (e.g., an equal probability of winning \$100 or \$101 indicates high certainty of getting some value around \$100), compared to gambles that have very different outcomes (e.g., equal probability of winning \$1 or \$200). Indeed, by varying the probabilities of the outcomes, Charpentier et al. (2018) and Iigaya et al. (2020) have demonstrated a positive relationship between the certainty of a positive outcome and information seeking behaviour–whether this relationship holds when the certainty is increased by manipulating the outcomes (and not the probabilities) is yet to be seen.

Ultimately, our results do not suggest that reward magnitude is unimportant in decision making in general–only that the extent to which it affects non-instrumental information seeking appears to be minimal. The present work represents more than a simple observation of the null effect of reward magnitude by offering a challenge to current information seeking theories and forcing the re-examination of how reward values should be considered in the decision making process.

Acknowledgements

The authors would like to thank Eleanor Psaila and Jake Embrey for conducting the experiments and analysing preliminary results. The project was supported by the Australian Research Council Grant DP190101076.

References

Aksoy, S., & Haralick, R. M. (2001). Feature normalization and likelihood-based similarity measures for image retrieval. *Pattern recognition letters*, *22*(5), 563–582.

Bennett, D., Bode, S., Brydevall, M., Warren, H., & Murawski, C. (2016). Intrinsic valuation of information in decision making under uncertainty. *PLoS computational biology*, *12*(7).

Charpentier, C. J., Bromberg-Martin, E. S., & Sharot, T. (2018). Valuation of knowledge and ignorance in mesolimbic reward circuitry. *Proceedings of the National Academy of Sciences*, *115*(31), E7255–E7264.

Ekenel, H. K., & Stiefelhagen, R. (2006). Analysis of local appearance-based face recognition: Effects of feature selection and feature normalization. In *2006 conference on computer vision and pattern recognition workshop (cvprw'06)* (pp. 34–34).

Embrey, J. R., Liew, S. X., Navarro, D., & Newell, B. R. (2020). Do you want to know a secret? the role of valence and delay in early informationpreference. In G. Denison, M. Mack, Y. Xu, & B. Armstrong (Eds.), *Proceedings of the 42nd annual conference of the cognitive science society* (p. 273-279). Austin, TX: Cognitive Science Society.

Iigaya, K., Hauser, T. U., Kurth-Nelson, Z., O'Doherty, J. P., Dayan, P., & Dolan, R. J. (2020). The value of what's to come: Neural mechanisms coupling prediction error and the utility of anticipation. *Science advances*, *6*(25), eaba3828.

Iigaya, K., Story, G. W., Kurth-Nelson, Z., Dolan, R. J., & Dayan, P. (2016). The modulation of savouring by prediction error and its effects on choice. *Elife*, *5*, e13747.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the american statistical association*, *90*(430), 773–795.

Sharot, T., & Sunstein, C. R. (2020). How people decide what they want to know. *Nature Human Behaviour*, 1–6.

Vasconcelos, M., Monteiro, T., & Kacelnik, A. (2015). Irrational choice and the value of information. *Scientific reports*, *5*(1), 1–12.

Zhu, J.-Q., Xiang, W., & Ludvig, E. A. (2017). Information seeking as chasing anticipated prediction errors. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. J. Davelaar (Eds.), *Proceedings of the 39th annual conference of the cognitive science society* (pp. 3658–3663). Austin, TX: Cognitive Science Society.