Considering what we know and what we don't know: Expectations and confidence guide value integration in value-based decision-making

Romy Frömer[a,*], Frederick Callaway[b,*], Thomas L. Griffiths[b], and Amitai Shenhav[a]

[a]Cognitive, Linguistic and Psychological Sciences, Brown University, Providence, RI
[b]Department of Psychology, Princeton University, Princeton, NJ

October 22, 2022

Abstract

When making decisions, we often have more information about some options than others. Previous work has shown that people are more likely to choose options that they look at more and those that they are more confident in. But should one always prefer options one knows more about? Intuition suggests not. Rather, how additional information impacts our preferences should depend critically on how valuable we expect the options to be. Here, we formalize this intuition in a Bayesian sequential sampling model where attention and confidence influence the precision of momentary evidence. Our model makes a key prediction: attention and confidence both increase choice probability for better-than-average options, and both decrease choice probability for worse-than-average options. We confirm this prediction in two experiments, in which we independently manipulate value and attention. Our results offer a novel perspective on prior work on the role of attention and confidence in decision-making, showing that people rely on contextual knowledge and uncertainty estimates to adaptively learn about their options and make better decisions.

There you are, at the overrun ice cream parlor. It is finally your turn, but despite having glared at the options for the last 10 minutes while waiting in line, you have not decided. The Stracciatella caught your eye early on, and you kept looking at it, imagining its creamy texture on your tongue occasionally interrupted by crispy chocolate chips. But as you moved down the line and caught glimpses of new flavors, you came to find it much less exciting than the pistachio blackberry swirl you just came to discover. You have never had that. Should you change your mind now? The Stracciatella would be good for sure, but the pistachio blackberry swirl might be amazing! Or, it could be a huge disappointment and make you sob for that Stracciatella. This example illustrates two ways in which decision-making is challenged by uncertainty. On one hand, information about our options becomes available asynchronously, for instance when we sequentially attend to our options. On the other hand, the quality of available information varies between options, for instance because we are more confident in the value of one compared to the other. How do decision-makers overcome these challenges?

Biases in attention have been ubiquitously linked to biases in choice; typically, options that are attended more are more likely to be chosen (Glickman et al., 2018; Usher et al., 2019; Zilker & Pachur, 2021). The mechanisms underlying attention's influence on choice have been formalized most prominently in the attentional drift diffusion model (aDDM; Krajbich et al., 2010). The aDDM extends sequential sampling models, in which evidence in favor of potential options is sampled repeatedly and accumulated over time (usually until some threshold is reached; Ratcliff et al., 2016; Bogacz et al., 2006; Usher & McClelland, 2001; Wang, 2012), with a parameter that discounts value information from unattended items. According to this model, choice biases arise from a multiplicative effect of value and attention on choice, such that in the appetitive domain, increased attention to an option increases the likelihood that it will be chosen, whereas in the aversive domain increased attention to an option decreases the likelihood that it will be chosen (Armel et al., 2008; Smith & Krajbich, 2019, but see Cavanagh et al., 2014; Westbrook et al., 2020; Kaanders et al., 2022).
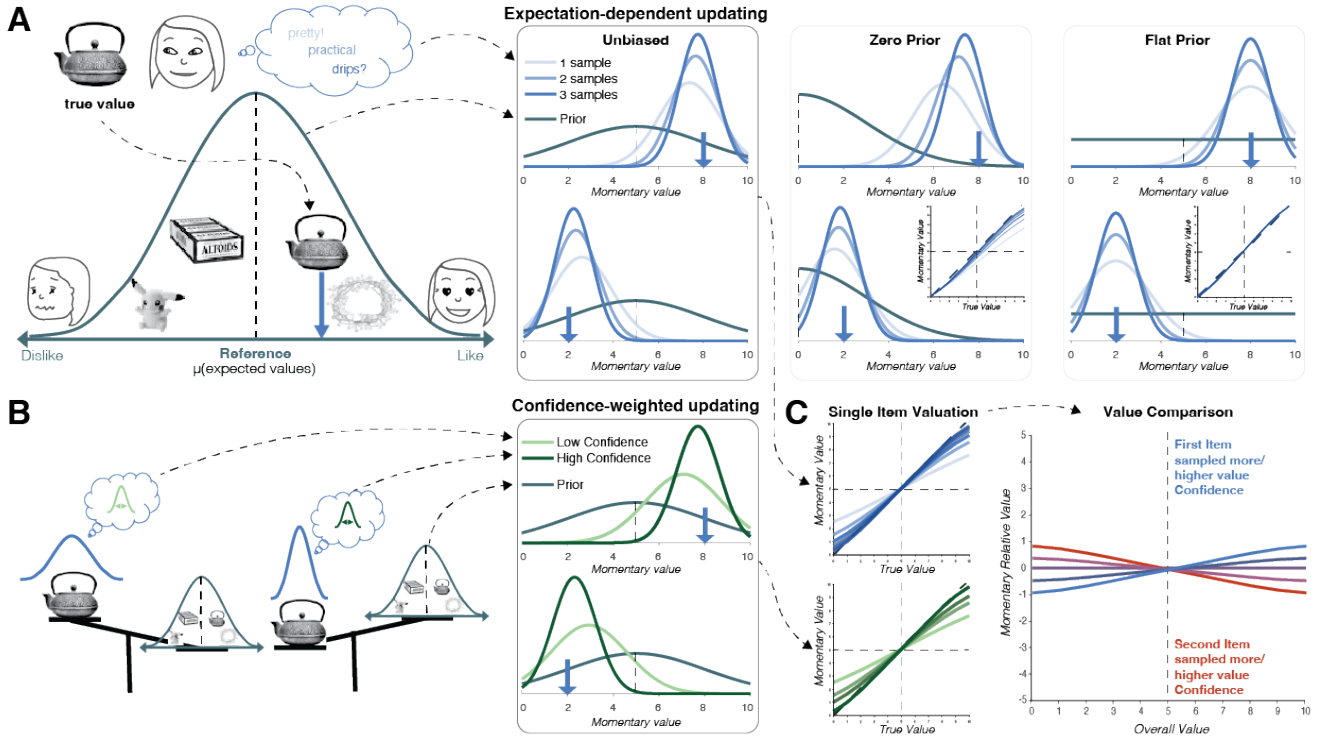
1

Figure 1: **Decision-making as sequential Bayesian updating predicts reference-dependent effects of attention and value confidence on choice. A** Prior expectations about the distribution of values in a given context (teal curve) influence how true item values (blue arrow) are inferred. Increasing amounts of information (increasingly blue curves) refine value estimates from the initial prior estimate towards the true value (blue arrow). With an unbiased prior, sampling typically increases the estimated value of better-than-average items (top), but decreases the estimate value for worse-than-average items (bottom). If, instead, the prior mean is zero (the minimum item value), sampling uniformly increases the momentary value estimate for all options. Alternatively, without expectations (flat prior), the estimated value will, on average, be exactly the true value, regardless of how much information is available. **B** Value confidence reflects the precision of the value representation underlying sampling. It thus determines the balance between samples and prior. When value confidence is high, samples are given greater weight and momentary value estimates are updated more. Similar to increasing the number of samples (e.g. through attention to an item), higher confidence leads to greater updating of the momentary value estimate away from the prior. Thus, like attention, value confidence has opposite effects for values above (top) and below (bottom) the mean. **C** Attention-guided sampling and value confidence similarly affect biases in momentary value estimates. Left panels show how momentary estimates systematically deviate from true values as a function of sampling and value confidence. Both decrease biases induced by the prior, hence for the representative (unbiased) prior highlighted in A and B, biases towards the average decrease with increased sampling and higher value confidence. During value comparison, asymmetries in sampling and value confidence induce systematic differences in momentary value estimates that translate into value-dependent choice biases. Increased sampling and higher confidence for one option over the other increase its choice probability when values are above the mean, but decrease it when values are below the mean.

Our confidence in the value of the available options can also impact our choices (Lee & Coricelli, 2020; Polanía et al., 2019; Quandt et al., 2021). Lower confidence in an option's value reflects greater variability in the value distribution underlying an options' value estimate (Lebreton et al., 2015), e.g. based on past experience (Quandt et al., 2021). This often renders an option less likely to be chosen (Lee & Coricelli, 2020). Furthermore, indicators of lower value confidence are associated with decreased decision consistency (De Martino

et al., 2013; Polanía et al., 2019), slower decision time, and decreased confidence in the choice that was ultimately made (Lee & Daunizeau, 2021).

This previous work has shown how biases in attention and lack of confidence can present a challenge for decision-makers. Here, we ask how people could flexibly and dynamically adapt to these challenges. We propose that both challenges can be understood as an imbalance or lack of information about the values of the choice options, and that

people compensate for this missing or imprecise information by drawing on expectations or contextual knowledge about likely option values (c.f. Polanía et al., 2019; Khaw et al., 2017). Formally, we propose a Bayesian extension of classic sequential sampling models in which attention and value confidence affect the quantity and quality of information about each item's value, which is integrated with a prior distribution to form posterior value estimates for each item.

As illustrated in Figure 1, our model makes the striking prediction that confidence and attention should both have opposite effects on choice probability (increasing vs. decreasing) when choosing among items better vs. worse than the average item observed over the course of the experiment. Items that are attended longer or for which value confidence is higher should be more likely to be chosen if their value is above the mean value, but less likely to be chosen if it is below the mean. We test and confirm this prediction in two experiments using a paradigm that allows us to manipulate attention independently of value. Taken together our results support the idea that decision-makers dynamically cope with uncertainty by considering both what they know and what they don't know.

Results

Bayesian evidence accumulation model

Following previous work (Tajima et al., 2016, 2019; Fudenberg et al., 2018; Callaway et al., 2021; Jang et al., 2021), we model value-based decision-making as an iterative process in which an agent forms posterior estimates of the value of each choice option based on a sequence of noisy value samples. Like the aDDM, this class of models assumes that the decision is based on a sequence of Gaussian samples whose mean depends on the value of each option in the choice set. However, while the aDDM assumes that the value samples are accumulated by simple addition, this newer class of models instead assumes that the samples are accumulated optimally, by Bayesian inference. Concretely, the agent estimates the true value of each item by inferring a posterior over the mean of the distribution from which the samples are drawn. Here, we present a simplified form of the model; see the Methods for the full specification.

At each time step, the agent attends to one of the available options, i, and receives a noisy signal of its subjective value,

$$x_t \sim \text{Normal}(u_{\text{true}}^{(i)}, 1/\tau^{(i)}). \tag{1}$$

Here, $u_{\text{true}}^{(i)}$ is the true value of the attended option and $\tau^{(i)}$ is the precision (inverse variance) of the samples. While previous work has assumed that this parameter is constant, we allow it to vary by item. $\tau^{(i)}$ thus captures differences in underlying confidence about the value of each option.

Given a sequence of samples and a prior $\text{Normal}(\mu_{\text{prior}}, 1/\lambda_{\text{prior}})$, the agent forms a posterior distribution over each item's value, $\text{Normal}(\mu_t^{(i)}, 1/\lambda_t^{(i)})$. Here, $\mu_t^{(i)}$ is the posterior mean (the estimated value) and $\lambda_t^{(i)}$ is the posterior precision (the certainty in that estimate). The posterior can be expressed in terms of the total number of samples taken for the item so far, $N_t^{(i)}$, and the mean of those samples, $\bar{x}_t^{(i)}$:

$$\mu_t^{(i)} = \mu_{\text{prior}} + \frac{N_t^{(i)} \tau_t^{(i)}}{\lambda_t^{(i)}} \left(\bar{x}_t^{(i)} - \mu_{\text{prior}} \right) \tag{2}$$
$$\lambda_t^{(i)} = \lambda_{\text{prior}} + N_t^{(i)} \tau^{(i)}$$

We have written the posterior in this way to emphasize two key properties of Bayesian updating. First, the impact of the samples on estimated value does not depend on their values *per se*, but rather on their values relative to one's expectations. In particular, the posterior mean begins at the prior mean and is shifted up or down depending on whether the sample mean is greater or less than the prior mean. This property, *expectation-dependent updating*, is illustrated in Figure 1A. Second, the degree to which the estimated value is shifted towards the sample mean (and, in expectation, the true value) depends on the precision of available information, where the total precision, $N_t^{(i)} \tau^{(i)}$, is composed of the quantity (number of samples, $N_t^{(i)}$) and quality (value confidence, $\tau^{(i)}$). This property, *precision-weighted updating* is illustrated in Figure 1B.

The most striking prediction of this model results from the interaction of these two properties, expectation dependence and precision weighting: the effect of both attention and confidence should have opposite effects on options whose value are above vs. below the reference (Figure 1C). In the following sections, we confirm this and other predictions of the model in two preferential choice experiments.

Considering what we know

We tested the proposal that people leverage knowledge about the distribution of values in the current choice context to dynamically compensate for uncertainty in two behavioral experiments in which participants made binary choices between consumer items. Most research on attention and choice has used gaze as a proxy for attention, allowing participants to freely fixate between the items in the choice set (Orquin & Mueller Loose, 2013). However, recent research has suggested that this fixation behavior is itself driven by estimated value and uncertainty (Song et al., 2019; Callaway et al., 2021; Jang et al., 2021; Li & Ma, 2021). Thus, to directly measure the causal influence of attention on choice, we instead manipulated fixation time to each item (cf. Armel et al., 2008; Tavares et al., 2017). Specifically, we controlled the order and duration that each item was displayed on the screen

(Figure 2A). In both studies, we constructed choice sets to vary in the relative and overall value of options based on participants' prior single item ratings that serve as estimates of the true values of the options. In Study 2 participants additionally provided confidence judgments for the initial value ratings (as in Lebreton et al., 2015), which we will examine below.

Our model makes predictions for how choices should deviate from utility maximization based on the true values if participants leveraged expectations about the values of options. As shown in Figure 1C, the momentary value of an item systematically deviates from its true value, depending on the number of samples taken for it and whether its true value is lower or higher than the reference (prior mean). When choosing among high-value options, more samples should lead to an advantage for that option, whereas when choosing among low-value options, more samples should lead to a disadvantage for that option. We would thus expect that, controlling for the relative value of options, choices would vary as a function of the interaction between the overall value of ones' options and the relative time spent sampling each of them (Figure 3A). In both studies, the variable of interest is therefore variability in behavior that is not explained by the value ratings. To test whether this residual variability followed the predictions made by our model, we regressed choices onto the relative presentation duration, overall value and their interaction, while controlling for relative value. We will first examine across both studies the predictions for attention effects on behavior and subsequently return to Study 2 to test the predictions for value confidence effects.

Participants' choices in both studies were sensitive to the values they had assigned in the earlier phase (Figure 2B). Participants were increasingly likely to choose either item as its value increased relative to the other item, $b_{S1} = 3.76, p < .001, b_{S2} = 3.39, p < .001$. Participants also made decisions faster when the value difference between items was larger, $b_{S1} = -0.22, p < .001, b_{S2} = -0.23, p < .001$ and, consistent with previous work (Frömer et al., 2019; Hunt et al., 2012; Smith & Krajbich, 2019), when the overall value of items was higher, $b_{S1} = -0.14, p < .001, b_{S2} = -0.17, p < .001$. Participants were also sensitive to the presentation order manipulation. In line with previous findings (Krajbich et al., 2010), all else being equal, participants were more likely to choose the first item they had seen, $b_{S1} = 0.10, p < .001, b_{S2} = 0.19, p < .001$, but this bias decreased for longer response times, $b_{S1} = -0.10, p < .001, b_{S2} = -0.23, p < .001$.

Controlling for all of these standard effects, we then tested our model's prediction about the three-way relationship between overall value, relative presentation duration, and choice: People should be more likely to choose items that are presented/seen longer when overall value is high, but less likely to choose them when overall value is low. Confirming this prediction, we found that presentation duration interacted with the overall value of the items $b_{S1} = 1.75, p < .001, b_{S2} =$
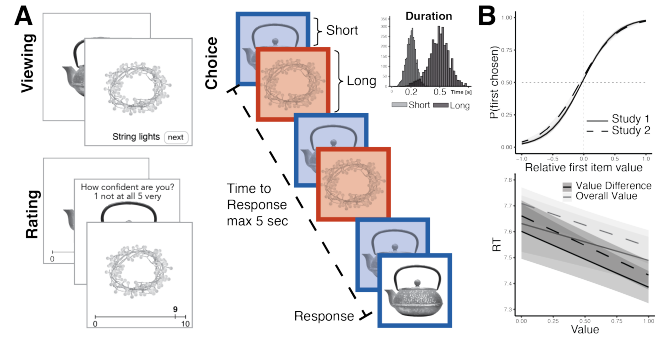

Figure 2: **Paradigm and behavioral results. A** In separate phases of the experiment, participants viewed consumer items (twice in Study 1, only), rated them, and (in Study 2, only) indicated the confidence in their ratings, and then chose among sets of two items, each. In the choice phase, items alternated on the screen until a choice was made or 5s elapsed. Frames around the options color-coded the corresponding response (e.g. left index for blue and right index for red). Presentation durations were manipulated between items, independently of presentation order, response hand, and value, and sampled from long (mean 500ms) and short (mean 200 ms) duration distributions, respectively. **B** Participants choices in both studies were sensitive to the relative value of options. Participants were faster when value difference was greater (black) and options were more valuable overall (grey).

$1.08, p = .016$, such that items that were presented longer were more likely chosen when overall value was high, but less likely chosen when overall value was low (Figure 3B).

We can further probe the degree to which participants calibrated their expectations to the distribution of values in their environment by testing the symmetry and strength of the interaction effect. How attention shapes value (and by extension, choice) depends on the mean of the prior (cf. Figure 1A). When people have calibrated their expectations to the value distribution, i.e., when the prior is unbiased, there should be no net effect of attention on choice. This is because attention has opposite effects on items below vs. above the prior mean, and these effects cancel out when the prior mean is the true mean item value. We implement two key alternative models with different priors: The first one assumes a prior of zero (i.e., that each presented item has minimal value), which results in a strictly positive effect of presentation duration on momentary value (for positive valance options, as in our experiment). Such a model was proposed by Jang et al. (2021) and is roughly consistent with the aDDM (Callaway et al., 2021). The second alternative assumes a flat prior, so that value estimates do not rely on expectations at all. As shown in Figure 3C and D, both models fail to capture the behavioral pattern: the zero-prior model fails to capture the opposite effects of presentation duration for high and low value items, and the flat-prior model fails to capture any interaction be-
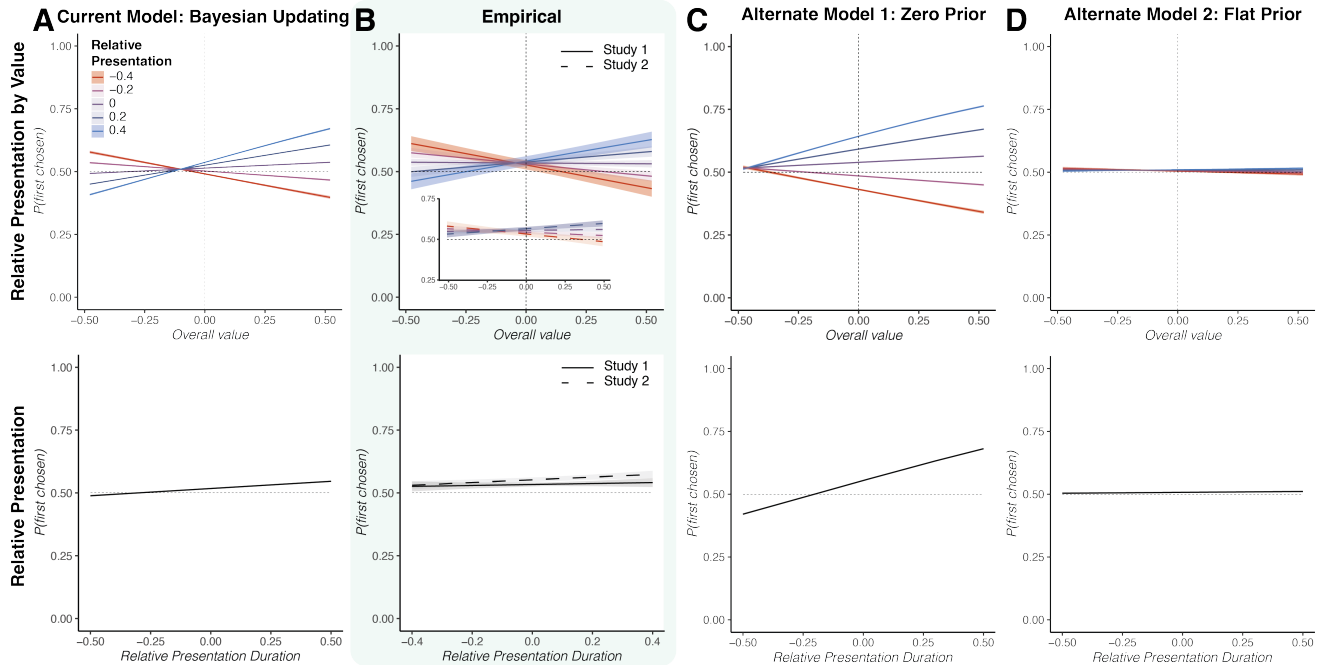
4

Figure 3: **People rely on contextual information about the expected distribution of values.** **A** According to our Bayesian model, a (relatively) unbiased prior should lead to value-dependent effects of relative presentation on choice, where choice probability increases with relative presentation for values above the mean, but decreases below the mean (top). An unbiased prior predicts no main effect of relative presentation on choice (bottom). **B** Empirical results in both studies are consistent with the predicted interaction (top) and the corresponding lack of a main effect. **C** A flat prior, with (near) uniform expectations of values does not produce the observed interaction. **D** A zero prior that overestimates the expected values also does not produce the observed interaction (top) and instead predicts a strong main effect of presentation duration on choice. In all plots, lines show predictions of fitted linear mixed effects models.

tween overall value and presentation duration.

Comparing Figure 3A and C reveals a clear behavioral signature of how well-calibrated our participants' expectations were. The cross-over point, at which the effect of relative presentation duration on choice is zero, corresponds to the prior mean. In the regression, this cross-over point is captured in the main effect of relative presentation duration on choice (a positive main effect if the cross-over point is below the true mean, and negative if above). Visually, we see that our participants' cross-over point was just slightly below zero, indicating a mean just slightly below the true empirical mean. Consistent with relatively unbiased prior means in both studies, we found no reliable main effects of presentation duration on choice $b_{S1} = 0.08, p = .634, b_{S2} = 0.22, p = .087$ (Figure 3C). With that said, the prior in Study 2 appeared more biased than in Study 1 as evinced by the larger estimate for the main effect of presentation duration (and the left-shifted cross-over point in Figure 4B). Additionally, unlike in Study 1, where the crossover point of the presentation-duration by overall value interaction is quite close to the empirical mean value, in Study 2 the crossover point occurred well below the empirical mean. This indicates that participants in Study 1

had a nearly unbiased prior mean, whereas the prior mean in Study 2 was more biased.

One plausible explanation for this descriptive difference is that participants in Study 1 had seen all items a second time before rating them and were given the opportunity to exclude unfamiliar items. We omitted this second round in Study 2 to increase the variance in confidence ratings and reduce the overall testing time. These design choices may have also prevented participants from fully updating a global default prior to the experimental context. However, our overall pattern of results shows that in both studies, participants clearly took information about the contextual distribution of values into account to some extent, and treated very low value items like aversive items (Shenhav et al., 2018; cf. Armel et al., 2008).

Taken together, our results are consistent with a Bayesian updating mechanism that relies on knowledge about the distribution of values in the task. As outlined earlier, this mechanism predicts the same effects for attention and confidence. We tested this prediction Using the value confidence ratings in Study 2.

Considering what we don't know

In Study 2, we tested how confidence in value ratings of the items in isolation affected choices among pairs of items. Previous work has shown that subjective confidence ratings provide a measure of the precision of one's value representation and the samples drawn from it (Lebreton et al., 2015; Quandt et al., 2021). Just like attention, confidence—or the lack thereof—should influence the extent to which momentary value estimates are biased towards the prior. Consequently confidence in item values should systematically bias choices between them.

A basic prediction of our model is that choices should be more consistent and faster when overall confidence in both items (set confidence) is higher (Figure 4A). We can compare these predictions to an alternative version of our model in which confidence determines the sample noise but not how the samples are integrated. This alternative model also predicts higher accuracy for higher set confidence, but at the cost of slower reaction times (see Figure 4C). Consistent with our model's prediction, Figure 4B shows that when their confidence in both items was higher, participants' choices were more consistent with their initial ratings (interaction of set confidence with relative value on choice $b = 0.48, p = .019$) and they responded faster (main effect of set confidence on RT $b = -0.02, p = .002$).

The key model predictions for relative confidence mirror those for relative presentation duration (Figure 4A, bottom; compare Figure 3A) and are not shared by the equal-weight updating model (Figure 4C, bottom), demonstrating that precision-weighted updating is a crucial component for this prediction. Specifically, we expect that relative confidence should interact with the overall value of options, such that high-confidence items are more likely to be chosen when their values are high and less likely to be chosen when their values are low. Our empirical data confirm this prediction (Figure 4B; interaction between relative confidence and overall value, $b = 0.63, p < .001$). Consistent with the biased prior mean mentioned earlier, this effect was not fully symmetric, so that we additionally observed a positive main effect of relative confidence on choice, $b = 0.11, p < .001$.

Our model and task also allow us to explore more subtle deviations from optimal Bayesian metacognition. Specifically, participants may be systematically over- or under-confident, treating samples as if they are more or less precise than they truly are. This would lead over-confident participants to make decisions that are faster and less consistent with their initial ratings, and under-confident participants to make decisions that are slower and more consistent with their initial ratings. As a proxy for overconfidence, we computed the average confidence across all items for each participant. When we include average confidence as a between-subject regressor in our analyses above, we observed the anticipated pattern. Participants with higher confidence tended to be faster

$b = -0.12, p = .060$, but less consistent $b = -1.13, p = .082$. While this pattern is consistent with our model prediction (Figure 5), the results do not reach the conventional statistical significance threshold of .05. Given that our sample is small and our measure of confidence bias is noisy, these results nevertheless encourage future work better suited to test this mechanism.

Discussion

When making decisions, we often need to overcome uncertainty due to attention-related delays in information or confidence-related imprecision of that information. In the present study we propose and test a Bayesian-updating account of value-based decision-making as a solution to these challenges. Under this theory, the value one assigns to one's options depends on two things: 1) one's expectations about the distribution of values in a given context, and 2) information acquired while considering those options. The relative weight of each component depends on the precision of that information. Crucially, the effect of acquiring more precise information about an option depends on how its true value compares to the expected value. If the true value is higher than expected, more information will tend to increase estimated value; but if the true value is lower than expected, more information will decrease estimated value. This theory thus makes the unique prediction that both increased attention to an option and greater confidence in its value will have opposite effects on choice probability for items that are below vs above the average value of options in the current context. This is exactly what we found.

Our results help shed new light on the role attention plays in decision-making. Previous work has proposed a multiplicative effect of attention and value on choice (Armel et al., 2008; Krajbich et al., 2010; Smith & Krajbich, 2019), such that when attended more, appetitive options would be more likely chosen, whereas aversive items would be less likely chosen. Indeed, only when appetitive and aversive items were included, has a full cross-over been previously demonstrated (Armel et al., 2008). Here we find the same effect, except that in our studies, the "bad" options were not truly aversive, but merely mundane—undesirable only when compared to the alternatives (Shenhav et al., 2018). Our finding that the effect of attention on choice *reverses* for mundane options therefore suggest that people can adapt their reference point, consistent with our Bayesian-updating account, as well as context-dependent range normalization (Khaw et al., 2017) and efficient coding (Polanía et al., 2019) accounts of valuation.

Our task allowed us to disentangle two effects that are typically conflated in studies of the influence of attention on value-based decision-making: the influence of attention on choice and the reverse influence of (beliefs about) choice on attention allocation. One finding that emerged as a result is
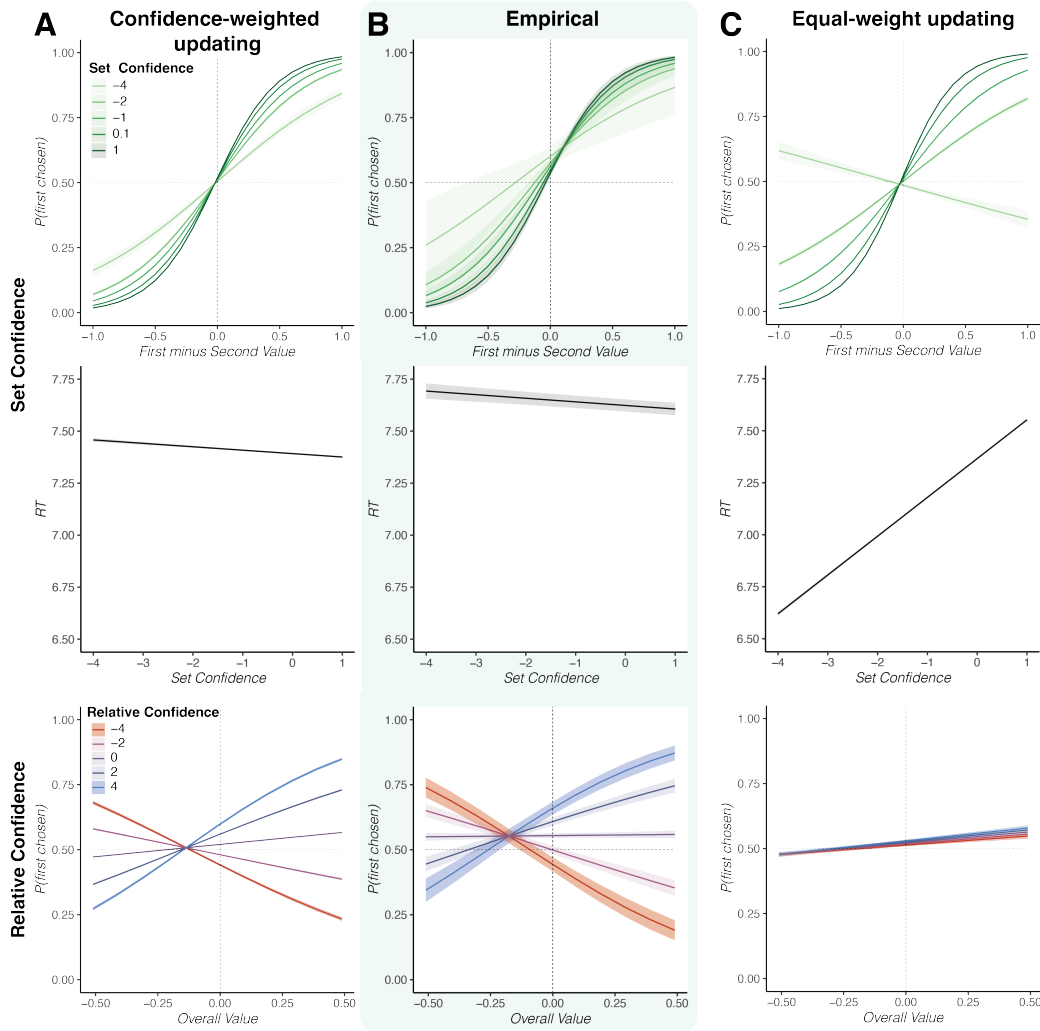
Figure 4: **People update proportional to their confidence in the values of their options. A** Our Bayesian updating model predicts more consistent and faster choices when choosing among options with higher value confidence. It also predicts an interaction of relative confidence with value, such that items with higher value confidence than the alternative should be more likely to be chosen when values are above the mean, but less likely when they are below the mean. **B** Empirical data in Study 2 are consistent with all these predictions. **C** A model in which the precision of the samples varies with value confidence, but all samples are weighted the same (average confidence) can only capture the increased choice consistency for higher value confidence, but neither of the other effects. It incorrectly predicts slower response times for higher set confidence and no effect of relative confidence.

that, unlike several past studies (Cavanagh et al., 2014; Krajbich et al., 2010; Westbrook et al., 2020), we did not find that attention exerted an independent (additive) influence on the likelihood of choosing one option or another, over and above its role in enhancing the influence of value on choice.

There are at least two potential reasons for this discrepancy. First, participants' expectations may have been biased in previous studies (Callaway et al., 2021; Jang et al., 2021). With few exceptions, in past studies only options that were positively evaluated during the initial rating or bidding phase were included in the subsequent choice sets. These excluded

items may have nevertheless exerted an influence on subsequent expectations about the distribution of option values (cf. Shenhav et al., 2018). This may have produced a prior closer to the mean-zero prior we introduced above, which predicts a main effect of attention on choice. However, this explanation may not be sufficient to explain the additive effect of attention on choice observed in previous studies, given that studies which include stimuli that were associated with negative feedback or mental effort (and were thus clearly aversive) still observed such an additive effect (Cavanagh et al., 2014; Westbrook et al., 2020).
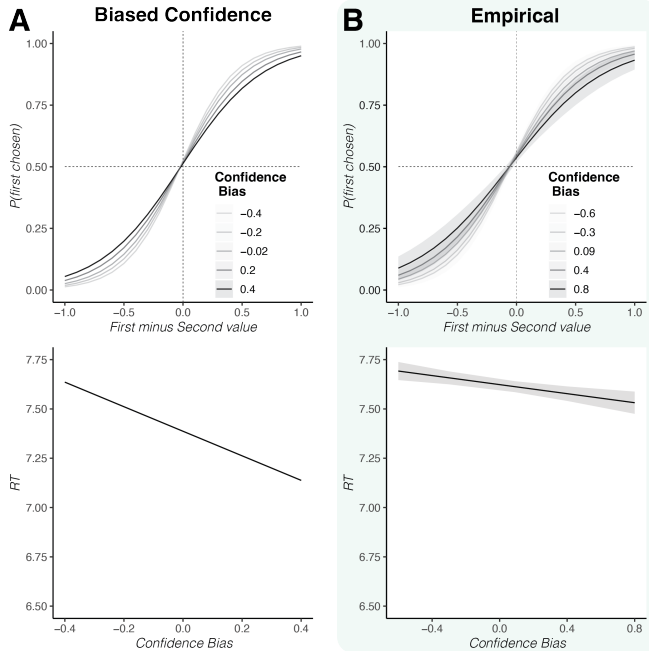
Figure 5: **Overconfidence leads to impulsive choice pattern. A** When making the model over- or underconfident, by varying the mapping between value confidence and the precision of the samples, it predicts that increasing overconfidence is associated with decreasing choice consistency and response times. **B** Our empirical results in Study 2 show the same pattern, where people with higher levels of confidence in all value ratings tend to respond less consistently and faster. These results, while consistent with our model are not significant at a conventional threshold.

The second potential explanation is that the relationship between attention and choice is not strictly unidirectional (Gluth et al., 2018, 2020; Callaway et al., 2021). In particular, participants in previous studies may have chosen to sample more information about the option they thought was best, producing a correlation between attention and choice through an entirely different causal mechanism (Shimojo et al., 2003; Kaanders et al., 2022). In our study, we intentionally broke the feedback-loop between the decision-process and gaze by experimentally manipulating presentation duration. Along these lines, it is notable that the only study that has shown a full crossover effect of attention and value on choice (Armel et al., 2008) also experimentally manipulated relative attention. In both our study and theirs, disabling active information-search may have eliminated the positive main-effect of gaze on choice. Future work will need to explicitly test this interpretation by comparing free-viewing and experimenter determined gaze in the same task.

Our value confidence results suggest that even when people can't actively seek information about specific options, they integrate novel information in a controlled way rather than simply uniformly accumulating information. Value confidence or equivalent measures are often thought of only as variables that impact the signal to noise ratio when sampling from representations with varying precision (Lebreton et al., 2015; Polanía et al., 2019). For example, one recent account suggests that value confidence determines the drift-rate of the decision-process within a DDM framework (Lee & Usher, 2021). Another account proposes that lower value confidence leads to increased deliberation time, but does not specify mechanistically how this would unfold (Lee & Daunizeau, 2021). Our model builds on this work by showing how people adaptively adjust their updating based on the precision of their value representations, resulting in both increased deliberation time and also increased reliance on prior knowledge.

Like attention, value confidence has been proposed to confer a choice benefit at least as long as options are appetitive (Lee & Coricelli, 2020; Li & Ma, 2021). On its face, this proposal resonates with ubiquitous uncertainty aversion, as well as the idea that confidence itself may be a value signal (Lebreton et al., 2015). Here we show that, as with attention, the influence of value confidence on choice critically depends on a decision-maker's expectations. This finding speaks against uncertainty aversion (or its converse, confidence bonus) as the main driver of choice variability associated with value confidence (Li & Ma, 2021). Rather it suggests a more nuanced role for value confidence in decision-making.

Our choice paradigm isolated one specific role for value confidence: weighing information in belief-updating. To the extent that the decision-maker thinks that the sampled information is precise (i.e., their confidence is high), they will adaptively reduce uncertainty by relying on samples versus expectations. This role parallels similar mechanisms in learning, where learners rely on response-based versus expectation-based predictions when they believe their evaluations are precise (Frömer et al., 2021). The subjectivity of confidence is an important aspect that sets our model apart from alternatives in which confidence merely reflects the signal-to-noise ratio (Lee & Usher, 2021). Our simulations and empirical results suggests that the extent to which one updates their beliefs is determined by their subjective *beliefs* about the precision of their samples, rather than merely the objective precision per se (see also Schiffer et al., 2017). When simulating mismatches between properties of the sampling distribution and subjective value confidence in our model, we produced impulsive, inconsistent behavior. We observed the same overall pattern in participants who were more confident overall. These findings add to a growing literature highlighting the importance of (well-calibrated) metacognition for adaptively regulating information-processing (Yeung & Summerfield, 2012; Frömer et al., 2021).

Value confidence likely plays a broader set of roles in decision-making beyond only regulating dynamic belief updating. For instance, previous work has shown that value confidence is used to regulate explore-exploit trade-offs in

learning contexts (Boldt et al., 2019). When attention is not experimentally manipulated, but can be freely allocated, we would expect that value confidence similarly provides inputs to (gaze-linked) information search. Importantly, the relationship between value confidence and gaze should then depend on the context (e.g. whether exploration is worthwhile; Wilson et al., 2014). It should also depend on the absolute level of confidence: if confidence is so low that no amount of sampling can increase certainty in the momentary value estimate to a reasonable level, participants should refrain from sampling an option, as the low expected information gain will not justify paying the cost of sampling. Unpacking this complementary role for value confidence in regulating information-processing in decision-making is an important future direction.

By controlling when and for how long items could be sampled through gaze, we eliminated the feedback-loop that has made it difficult to interpret correlations between value, attention, and choice (discussed above). The flipside of this is that we still need to test predictions of models that both allow for free gaze allocation and incorporate confidence-dependent updating. Indeed, confidence could have complex, time-varying effects on attention allocation. Early in the decision, attending to a high-confidence item produces stronger and thus more valuable information. Later in the decision, however, the value of the high-confidence item will have a very precise estimate, making additional attention to it superfluous. By isolating the downstream effects of confidence on choice, we have provided the foundation for understanding the full cyclic process.

Another open question is whether goal-directed information search is limited to simply which option is considered, or in fact extends to what type of information is considered about a given option. Most sequential sampling models assume that samples are drawn independently from an underlying distribution, but this need not be so. In the context of memory retrieval, during free recall, people are more likely to sequentially sample items that share features (e.g. category) than dissimilar items (Bousfield & Cohen, 1955), perhaps due to spreading of activation that facilitates the retrieval of related items. To account for these findings, Zhu et al. (2018) propose a sequential sampling model in which successive samples are drawn from a Markov chain, and are hence autocorrelated. Based on this work, we would expect autocorrelation in value-based sampling as well. For example, when sampling one positive feature about the tea pot in Figure 1 (nice color), one would be more likely to sample other related positive features about it (great shape), rather than a less related feature or a negative one (might drip).

Importantly, this autocorrelation might be biased by goals and momentary beliefs, just like gaze (Gluth et al., 2020; Callaway et al., 2021; Jang et al., 2021) and explicit information gathering (Hunt et al., 2016). Reports of confirmation bias abound (Kaanders et al., 2022; Sharot & Garrett, 2016; Talluri et al., 2018), and classic work on framing (Morewedge & Kahneman, 2010), as well as recent work on choice goals' influence on decision making (Frömer et al., 2019; Sepulveda et al., 2020), show that people's prior beliefs and the type of question they are trying to answer shape the degree to which they integrate information and seek it to begin with. How these mechanisms shape value-based choices remains to be understood.

One canonical finding that might be explained by biased memory sampling is the speeding of choices among options that are more rather than less congruent with one's choice goal (Frömer et al., 2019; Hunt et al., 2012; Sepulveda et al., 2020; Smith & Krajbich, 2019). Different proposals have been put forward to explain the overall value effect on choice. One account explains the effect through attentional biases (Sepulveda et al., 2020; Smith & Krajbich, 2019), whereas in other accounts these effects emerge from the dynamics of the decision-process without considering gaze variability (Frömer et al., 2019; Hunt et al., 2012). In a reference-dependent value coding scheme as demonstrated here, none of the current models are able to account for the overall value effect on decision time. In fact, our model can only produce it due to its slightly biased prior and the simulated effects are much weaker than those found in the data. This tension between predicted and observed findings might suggest that caution is warranted in assuming a role for reference-dependence in this process. However, given the convergent evidence of reference-dependence from our remaining results and other lines of work (Shenhav et al., 2018; Khaw et al., 2017; Polanía et al., 2019), a more plausible alternative is that goal-congruent information is sampled with greater ease, at a faster rate than goal-incongruent information, in line with confirmation and positive evidence biases in information sampling (Hunt et al., 2016; Kaanders et al., 2022; Talluri et al., 2018; Sharot & Garrett, 2016). Future work will need to explicitly manipulate prior expectations and choice goals to arbitrate between different accounts.

Our findings have implications for interpretations of neural correlates of value in value-based choice. Traditionally, neuroeconomics has focused on valuation and evidence accumulation when interpreting neural correlates of value, for instance in dorsal anterior cingulate cortex (Hare et al., 2011; Pisauro et al., 2017). However, these regions have also been associated with goal-directed regulation of information gain (McGuire et al., 2014; O'Reilly et al., 2013), and cognitive control more generally (Shenhav et al., 2013). Our results contribute to a growing consensus that value-based choice draws on those regulatory mechanisms, suggesting that the observed activity in dACC could reflect information regulation rather than accumulated evidence per se (Hunt et al., 2018; Hunt, 2021; Hunt et al., 2021; Kaanders et al., 2020; Monosov, 2017).

By characterizing the specific quantities associated with information regulation in the service of decision making (e.g.

value confidence, estimate precision, changes in choice certainty), our model can inform specific predictions for neural activity and crucially how it unfolds over time (c.f. Frömer et al., 2021; McGuire et al., 2014; O'Reilly et al., 2013). Testing these predictions will allow us to re-evaluate and extend beyond previously identified correlates of evidence accumulation in value-based choice (e.g., Frömer & Shenhav, 2019; Pisauro et al., 2017; Polania et al., 2014).

A central motivation for the decision sciences is to understand why decision-making is so often flawed or irrational and to use that knowledge to improve decision-making. Viewing decision-making as an active process, involving information search and regulation of information-processing shaped by goals and metacognition, provides us with a whole new set of levers to achieve this goal. Perhaps if we understand how these control mechanisms work, we can help decision-makers and learners alike by teaching them to consider and control what they know and what they don't know.

Methods

Experiment

Participants

Participants were recruited from Brown University and the general community. Thirty participants (19 female) with an average age of 19.6 (SD = 1.8) took part in Study 1. Thirty one participants (23 female) with an average age of 21.1 (SD = 3.5) took part in Study 2. Participants gave informed consent and were remunerated with 10 USD per hour or course credits. The studies were approved by Brown University's institutional review board.

Task and Procedure

In Study 1, the experiment consisted of three phases: 1) item familiarization, 2) item rating, and 3) choice. During item familiarization, participants viewed consumer items on the screen individually in a self-paced manner. Items were presented in grey scale and short labels were displayed below the items. Following a first round of viewing all items, participants were presented with each item a second time without labels and asked to indicate items that they could not identify without a label present. These items were removed from subsequent phases of the experiment. In the subsequent rating phase participants were asked to rate how much they liked the remaining items individually on a scale from 0 (not at all) to 10 (a great deal). They were encouraged to use the entire range of values rather than just the extremes. Based on the individual ratings, we constructed personalized choice sets that varied in the relative and overall values of options (Shenhav & Buckner, 2014; Shenhav & Karmarkar, 2019; Frömer et al., 2019; Shenhav et al., 2018). Depending on the range of

values participants provided and the number of items they excluded during familiarization, the number of choice sets we could generate varied between 128 and 240 (Median = 228, M = 222, SD = 24).

In the choice phase, participants viewed pairs of options alternating centrally on the screen inside frames that color coded the corresponding response hands (index fingers placed on A and J keys on a standard keyboard). We manipulated the relative presentation duration of items so that one item was always presented longer than the other. While items alternated on the screen, specific durations on each turn were drawn from different distributions for long (M = 500 ms, SD = 100 ms) and short presentations (M = 200, SD = 50); these distributions were informed by previous work (Frömer & Shenhav, 2019). Presentation duration, response hand, and order of presentation varied independently. Items alternated until participants made a choice or 5s had elapsed. After making their choice, participants were presented with the item they chose for one second and allowed to reverse their choice within 750 ms of feedback presentation (5.6% of all trials). Participants performed multiple rounds of practice with empty frames only, letters, and with practice items, to learn the button mapping and get comfortable with the task.

Study 2 differed from Study 1 in three ways: 1) participants performed only the first round of familiarization and no items were excluded, 2) in the rating phase individual item ratings were immediately followed by a prompt to indicate the confidence in that rating, and 3) item ratings were repeated after the choice phase. We removed the second round of familiarization to reduce overall testing time and to include items with very low value confidence. Participants rated their confidence in the value ratings on a scale from 1 (not at all) to 5 (very) by pressing the corresponding number key on a keyboard. The second round of value ratings following the choice phase did not include confidence ratings. Participants performed 240 choice trials unless their rating distributions afforded less (Median = 240, Mean = 236, SD = 7). On average they indicated they had made an error 5.9% of all trials.

Analyses

Our primary interest was participants' behavior during the choice phase. We analyzed the probability of choosing the first item using generalized mixed effects models with a binomial link function, and RTs using linear mixed effects models using lme4 (Bates, Maechler, et al., 2015) in R (Version 3.6.1). We modeled random intercepts for participants and included random slopes for predictors if supported by the data (Bates, Kliegl, et al., 2015; Matuschek et al., 2017). We performed model selection using the buildmer package. Choices that were made before the second item was seen were excluded from all analyses (0.7% in Study 1, 0.3% in Study 2). The remaining choices were analyzed with relative value (first item value minus second item value), overall value (average

value of both items in a set), relative presentation duration (proportion of time the first item was on the screen), and RT as an index of total decision time. To test how relative presentation duration influences values' relationship with choice behavior, we modeled interactions between relative presentation duration and value regressors. We analyzed response times with value difference (absolute difference between item values) and overall value. In Study 2, we further included three variables to assess the impact of value confidence on choice: relative confidence (confidence in the first item value minus confidence in the second item value), set confidence (mean confidence in both items, centered within participants), and confidence bias as a between subject regressor (average confidence of a participant across all items). To test how value confidence impacts values' relationship with choice behavior, we modeled interactions between confidence and value regressors.

Model

Here we give further details on the implementation of the model.

Attention

In the simplified model presented in the main text, we assumed that attention controls which item is sampled at each time step. The full model, however, allows for partial attention to the non-presented item. Following Jang et al. (2021), we implement partial attention as reduced precision. Thus, at each time step, the agent receives one signal for each item,

$$x_t^{(i)} \sim \text{Normal}\left(u_{\text{true}}^{(i)}, \frac{1}{\theta_t^{(i)}\tau^{(i)}}\right), \qquad (3)$$

where $\theta_t^{(i)}$ captures the degree of attention to item i at time t. It is defined:

$$\theta_t^{(i)} = \begin{cases} 1 & \text{if } i \text{ is attended at time } t \\ \theta & \text{if } i \text{ has been attended at some } t' < t \\ 0 & \text{if } i \text{ has not been attended before } t \end{cases} \qquad (4)$$

That is, the precision of samples for the unattended item is discounted by a multiplicative factor, θ, a free parameter between 0 and 1. On the first presentation, the unattended item has zero precision as it has not been seen yet, and so samples of its value cannot possibly be drawn.

To account for partial attention in the posterior (Equation 2), we generalize $N_t^{(i)}$ (previously the number of samples) to a continuous quantity capturing the amount of attention paid to the item so far,

$$N_t^{(i)} = \sum_{j=1}^{t} \theta_t^{(i)}. \qquad (5)$$

Similarly, $\bar{x}_t^{(i)}$ becomes an attention-weighted mean of the samples,

$$\bar{x}_t^{(i)} = \frac{1}{N_t^{(i)}} \sum_{j=1}^{t} \theta_t^{(i)} x_t^{(i)} \qquad (6)$$

Confidence

Value confidence is captured in the $\tau^{(i)}$ term. More precisely, $\tau^{(i)}$ is the precision of samples for an item when it is attended. Here, we refer to this parameter as "baseline precision" to distinguish it from confidence ratings (which are only indirectly related to precision, see below). In the first study, we did not collect confidence measures, so we assume that all items have the same baseline precision, a free parameter. In the second study, we assume that baseline precision varies by item, and that this variance is related to the rating confidence judgements. For simplicity, we assume that an item's baseline precision is an affine function of its confidence rating,

$$\tau^{(i)} = \beta_0 + \beta_1 \cdot \text{conf}^{(i)}, \qquad (7)$$

where β_0 and β_1 are free parameters.

In the main model, we assume that both the samples themselves and the posterior estimate (given the samples) are constructed using a single $\tau^{(i)}$ parameter, as in optimal Bayesian inference. However, we also consider the possibility that the posterior is formed assuming an incorrect precision, replacing $\tau^{(i)}$ with $\tilde{\tau}^{(i)}$ in Equation 2. We consider two forms of incorrect precision. In the "average confidence" model, the agent treats all items as though they have the same sampling precision:

$$\tilde{\tau}^{(i)} = \beta_0 + \beta_1 \cdot \text{mean}(\text{conf}). \qquad (8)$$

In the over/under-confidence model, the agent assumes that sampling precision is systematically higher or lower than it really is:

$$\tilde{\tau} = \beta_0 + \gamma\beta_1 \cdot \text{conf}, \qquad (9)$$

where $\gamma > 1$ corresponds to overconfidence and $\gamma < 1$ corresponds to underconfidence.

Critically, in both of these models, the true noisiness of the samples (Equation 3) remains unchanged. It is only the posterior (Equation 2) that uses the biased $\tilde{\tau}^{(i)}$ term.

Prior

Following Callaway et al. (2021), we assume that the prior takes the form of a Gaussian distribution with a mean and standard deviation related to the actual empirical distribution of values as follows:

$$\mu_{\text{prior}} = \alpha \cdot \text{mean}(\text{ratings})$$
$$\lambda_{\text{prior}} = \text{std}(\text{ratings})^{-2}. \qquad (10)$$

Here, $\text{std}(\text{ratings})$ denotes the standard deviation of the value ratings for all items used in the choice phase. The α parameter

allows us to capture systematic biases in the prior, which may result from incomplete adaptation to the experimental context. In the full model, α is a free parameter. In the zero-prior model, α is fixed to zero. Finally, in the flat-prior model, we set λ_{prior} to 10^{-6}, effectively removing all prior information.

Stopping rule

Intuitively, one should stop sampling (and make a decision) when the cost of additional samples exceeds the expected increase in decision quality from collecting those samples. Previous work has used dynamic programming to identify the optimal stopping rule (Drugowitsch et al., 2012; Tajima et al., 2016). However, the exogenous attention manipulation complicates this approach, and the exact shape of the stopping rule is not of critical interest for the current study. Thus, we instead employ a simple approximation to the optimal stopping rule. The approximation has two parts: First, rather than predicting the specific sequence of future presentation durations, we assume that all future samples will be split between the items according to the ratio of their average presentation durations. Second, we estimate the value of additional sampling using an adaptation of the Directed Cognition model proposed by Gabaix & Laibson (2005). Concretely, we compute the expected increase in decision quality (the value of the chosen option) for different ammounts of additional sampling, and substract the cost of that additional sampling. If there is any amount of sampling for which this value is positive, the model continues sampling. See the Appendix for further details.

Acknowledgements

References

Armel, K. C., Beaumel, A., & Rangel, A. (2008). Biasing simple choices by manipulating relative visual attention. *Judgment and Decision Making*, *3*(5), 396-403.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv preprint arXiv:1506.04967*.

Bates, D., Maechler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1-48.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*(4), 700–765.

Boldt, A., Blundell, C., & De Martino, B. (2019). Confidence modulates exploration and exploitation in value-based learning. *Neurosci Conscious*, *2019*(1), niz004.

Bousfield, W. A., & Cohen, B. H. (1955). The occurrence of clustering in the recall of randomly arranged words of different frequencies-of-usage. *Journal of General Psychology*, *52*, 83.

Callaway, F., Rangel, A., & Griffiths, T. L. (2021, March). Fixation patterns in simple choice reflect optimal information sampling. *PLOS Computational Biology*, *17*(3), e1008863.

Cavanagh, J. F., Wiecki, T. V., Kochar, A., & Frank, M. J. (2014). Eye tracking and pupillometry are indicators of dissociable latent decision processes. *J Exp Psychol Gen*, *143*(4), 1476-88.

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

Drugowitsch, J., Moreno-Bote, R., Churchland, A. K., Shadlen, M. N., & Pouget, A. (2012). The cost of accumulating evidence in perceptual decision making. *Journal of Neuroscience*, *32*(11), 3612–3628.

Frömer, R., Dean Wolf, C. K., & Shenhav, A. (2019). Goal congruency dominates reward value in accounting for behavioral and neural correlates of value-based decision-making. *Nature Communications*, *10*(1), 4926.

Frömer, R., & Shenhav, A. (2019). Spatiotemporally distinct neural mechanisms underlie our reactions to and comparison between value-based options. *bioRxiv*, 609198.

Frömer, R., Nassar, M. R., Bruckner, R., Stürmer, B., Sommer, W., & Yeung, N. (2021). Response-based outcome predictions and confidence regulate feedback processing and learning [Journal Article]. *eLife*, *10*, e62825.

Fudenberg, D., Strack, P., & Strzalecki, T. (2018). Speed, accuracy, and the optimal timing of choices. *American Economic Review*, *108*(12), 3651–3684.

Gabaix, X., & Laibson, D. (2005). Bounded Rationality and Directed Cognition. *Working Paper*.

Glickman, M., Tsetsos, K., & Usher, M. (2018). Attentional selection mediates framing and risk-bias effects. *Psychological Science*, *29*(12), 956797618803643.

Gluth, S., Kern, N., Kortmann, M., & Vitali, C. L. (2020). Value-based attention but not divisive normalization influences decisions with multiple alternatives. *Nature Human Behaviour*, *4*(6), 634-645.

Gluth, S., Spektor, M. S., & Rieskamp, J. (2018). Value-based attentional capture affects multi-alternative decision making. *Elife*, *7*, e39659.

Hare, T. A., Schultz, W., Camerer, C. F., O'Doherty, J. P., & Rangel, A. (2011). Transformation of stimulus value signals into motor commands during simple choice. *Proc Natl Acad Sci U S A*, *108*(44), 18120-5.

Hunt, L. T. (2021). Frontal circuit specialisations for decision making [Journal Article]. *Eur J Neurosci*, *53*(11), 3654-3671.

Hunt, L. T., Daw, N. D., Kaanders, P., MacIver, M. A., Mugan, U., Procyk, E., ... Kolling, N. (2021). Formalizing planning and information search in naturalistic decision-making [Journal Article]. *Nature Neuroscience*.

Hunt, L. T., Kolling, N., Soltani, A., Woolrich, M. W., Rushworth, M. F., & Behrens, T. E. (2012). Mechanisms underlying cortical activity during value-guided choice. *Nat Neurosci*, *15*(3), 470-6, S1-3.

Hunt, L. T., Malalasekera, W. M. N., de Berker, A. O., Miranda, B., Farmer, S. F., Behrens, T. E. J., & Kennerley, S. W. (2018). Triple dissociation of attention and decision computations across prefrontal cortex. *Nat Neurosci*, *21*(10), 1471-1481.

Hunt, L. T., Rutledge, R. B., Malalasekera, W. M., Kennerley, S. W., & Dolan, R. J. (2016). Approach-induced biases in human information sampling. *PLoS Biol*, *14*(11), e2000638.

Jang, A. I., Sharma, R., & Drugowitsch, J. (2021, March). Optimal policy for attention-modulated decisions explains human fixation behavior. *eLife*, *10*, e63436.

Kaanders, P., Nili, H., O'Reilly, J. X., & Hunt, L. T. (2020). Medial frontal cortex activity predicts information sampling in economic choice. *bioRxiv*, 2020.11.24.395814.

Kaanders, P., Sepulveda, P., Folke, T., Ortoleva, P., & De Martino, B. (2022). Humans actively sample evidence to support prior beliefs [Journal Article]. *Elife*, *11*, e71768. doi: 10.7554/eLife.71768

Khaw, M. W., Glimcher, P. W., & Louie, K. (2017). Normalized value coding explains dynamic adaptation in the human valuation process. *Proceedings of the National Academy of Sciences*, *114*(48), 12696-12701.

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nat Neurosci*, *13*(10), 1292-8.

Lebreton, M., Abitbol, R., Daunizeau, J., & Pessiglione, M. (2015). Automatic integration of confidence in the brain valuation signal. *Nat Neurosci*, *18*(8), 1159-67.

Lee, D. G., & Coricelli, G. (2020). An empirical test of the role of value certainty in decision making. *Frontiers in Psychology*, *11*(2485).

Lee, D. G., & Daunizeau, J. (2021). Trading mental effort for confidence in the metacognitive control of value-based decision-making [Journal Article]. *Elife*, *10*, e63282. doi: 10.7554/eLife.63282

Lee, D. G., & Usher, M. (2021). Value certainty in drift-diffusion models of preferential choice. *Psychological Review*. doi: 10.1037/rev0000329

Li, Z. W., & Ma, W. J. (2021). An uncertainty-based model of the effects of fixation on choice. *PLoS Comput Biol*, *17*(8), e1009190. doi: 10.1371/journal.pcbi.1009190

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing type i error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305-315.

McGuire, J. T., Nassar, M. R., Gold, J. I., & Kable, J. W. (2014). Functionally dissociable influences on learning rate in a dynamic environment. *Neuron*, *84*(4), 870-81.

Monosov, I. E. (2017). Anterior cingulate is a source of valence-specific information about value and uncertainty. *Nature Communications*, *8*(1), 134.

Morewedge, C. K., & Kahneman, D. (2010). Associative processes in intuitive judgment. *Trends Cogn Sci*, *14*(10), 435-40. doi: 10.1016/j.tics.2010.07.004

Nadarajah, S., & Kotz, S. (2008). Exact distribution of the max/min of two gaussian random variables. *IEEE Transactions on very large scale integration (VLSI) systems*, *16*(2), 210–212.

O'Reilly, J. X., Schuffelgen, U., Cuell, S. F., Behrens, T. E., Mars, R. B., & Rushworth, M. F. (2013). Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proc Natl Acad Sci U S A*, *110*(38), E3660-9.

Orquin, J. L., & Mueller Loose, S. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica*, *144*(1), 190–206.

Pisauro, M. A., Fouragnan, E., Retzler, C., & Philiastides, M. G. (2017). Neural correlates of evidence accumulation during value-based decisions revealed via simultaneous eeg-fmri. *Nat Commun*, *8*, 15808.

Polania, R., Krajbich, I., Grueschow, M., & Ruff, C. C. (2014). Neural oscillations and synchronization differentially support evidence accumulation in perceptual and value-based decision making. *Neuron*, *82*(3), 709-20.

Polanía, R., Woodford, M., & Ruff, C. C. (2019). Efficient coding of subjective value. *Nature Neuroscience*, *22*(1), 134-142.

Quandt, J., Figner, B., Holland, R. W., & Veling, H. (2021). Confidence in evaluations and value-based decisions reflects variation in experienced values [Journal Article]. *J Exp Psychol Gen*. doi: 10.1037/xge0001102

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends Cogn Sci*, *20*(4), 260-81.

Schiffer, A. M., Siletti, K., Waszak, F., & Yeung, N. (2017). Adaptive behaviour and feedback processing integrate experience and instruction in reinforcement learning. *Neuroimage*, *146*, 626-641.

Sepulveda, P., Usher, M., Davies, N., Benson, A., Ortoleva, P., & De Martino, B. (2020). Visual attention modulates the integration of goal-relevant evidence and not value. *bioRxiv*, 2020.04.14.031971.

Sharot, T., & Garrett, N. (2016). Forming beliefs: Why valence matters [Journal Article]. *Trends Cogn Sci*, *20*(1), 25-33.

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron*, *79*(2), 217-40.

Shenhav, A., & Buckner, R. L. (2014). Neural correlates of dueling affective reactions to win–win choices. *Proceedings of the National Academy of Sciences*, *111*(30), 10978-10983.

Shenhav, A., Dean Wolf, C. K., & Karmarkar, U. R. (2018). The evil of banality: When choosing between the mundane feels like choosing between the worst. *J Exp Psychol Gen*, *147*(12), 1892-1904.

Shenhav, A., & Karmarkar, U. R. (2019). Dissociable components of the reward circuit are involved in appraisal versus choice. *Sci Rep*, *9*(1), 1958.

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nat Neurosci*, *6*(12), 1317-22.

Smith, S. M., & Krajbich, I. (2019). Gaze amplifies value in decision making. *Psychological Science*, *30*(1), 116-128.

Song, M., Wang, X., Zhang, H., & Li, J. (2019). Proactive information sampling in value-based decision-making: Deciding when and where to saccade. *Frontiers in Human Neuroscience*, *13*(February), 1–10.

Tajima, S., Drugowitsch, J., Patel, N., & Pouget, A. (2019). Optimal policy for multi-alternative decisions. *Nat Neurosci*, *22*(9), 1503-1511.

Tajima, S., Drugowitsch, J., & Pouget, A. (2016). Optimal policy for value-based decision-making. *Nat Commun*, *7*, 12400.

Talluri, B. C., Urai, A. E., Tsetsos, K., Usher, M., & Donner, T. H. (2018). Confirmation bias through selective overweighting of choice-consistent evidence. *Curr Biol*, *28*(19), 3128-3135 e8.

Tavares, G., Perona, P., & Rangel, A. (2017). The attentional Drift Diffusion Model of simple perceptual decision-making. *Frontiers in Neuroscience*, *11*(AUG), 1–16.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*(3), 550-592.

Usher, M., Tsetsos, K., Glickman, M., & Chater, N. (2019). Selective integration: An attentional theory of choice biases and adaptive choice. *Current Directions in Psychological Science*, *28*(6), 552-559.

Wang, X. J. (2012). Neural dynamics and circuit mechanisms of decision-making. *Curr Opin Neurobiol*, *22*(6), 1039-46.

Westbrook, A., van den Bosch, R., Maatta, J. I., Hofmans, L., Papadopetraki, D., Cools, R., & Frank, M. J. (2020). Dopamine promotes cognitive effort by biasing the benefits versus costs of cognitive work. *Science*, *367*(6484), 1362-1366.

Wilson, R. C., Geana, A., White, J. M., a Ludvig, E., & 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.

Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: confidence and error monitoring. *Philos Trans R Soc Lond B Biol Sci*, *367*(1594), 1310-21.

Zhu, J., Sanborn, A., & Chater, N. (2018). Mental Sampling in Multimodal Representations. In *Advances in Neural Information Processing Systems* (Vol. 31). Curran Associates, Inc.

Zilker, V., & Pachur, T. (2021). Nonlinear probability weighting can reflect attentional biases in sequential sampling. *Psychological Review*.

Stopping rule

As briefly described in the main text, we use a two-part approximation to the optimal stopping rule. Our method approximates the information content of future samples by assuming that both items will be presented in proportion to their average presentation times; it then estimates the value of additional sampling by hypothetically committing to different amounts of sampling. We provide further details on each component below.

Approximating future sample precisions A truly optimal stopping policy would take into account the specific sequence of future presentations, in particular, using the elapsed time in the current presentation to predict the switch and therefore the precision of upcoming samples. Rather than model this complex (and perhaps psychologically implausible) process, we instead use the proportion of presentation time on each item. That is, we use a mean-field approximation with respect to time. Concretely, we approximate the attentional downweighting for each item as

$$\bar{\theta}^{(1)} = \frac{\bar{t}_1 + \bar{t}_2 \theta}{\bar{t}_1 + \bar{t}_2}, \tag{11}$$

where $\bar{t}_1$ and $\bar{t}_2$ are the means of the presentation duration distributions for the two items (200ms and 500ms). $\bar{\theta}^{(1)}$ is defined analogously.

Value of additional sampling Given these precisions, how should one decide whether to make a decision or keep sampling? We approximate the optimal stopping rule using an adaptation of the Directed Cognition model proposed by Gabaix & Laibson (2005). The intuition for the approximation comes in three steps.

First, we can quantify the *value of information* (VOI) for taking one additional sample as the expected increase in the utility of the item that would be chosen with additional information vs. immediately:

$$\text{VOI}_t(1) = \text{E}\left[u_{\text{choice}_{t+1}} - u_{\text{choice}_t}\right]. \tag{12}$$

We derive a closed-form expression for the VOI below.

Second, if the expected increase in decision quality for taking one more sample is greater than the sample cost, then clearly one should take (at least) one more sample. That is, one should take another sample if

$$\text{VOI}_t(1) < c \tag{13}$$

where c is the cost per sample, a free parameter of the model.

Third, note that $\text{VOI}_t(1)$ underestimates the true value of sampling. This is because it assumes that one has to make a decision after taking one additional sample when one could in fact take more samples. In some cases, it may require multiple samples to change one's mind; in these cases, sampling once

has minimal value, while sampling several times could have substantial value. To mitigate this problem, we can hypothetically commit to taking N samples. If the VOI for N more samples is greater than the cost of those samples (for any N), then it would be better to commit to taking N more samples than to make a decision now. However, actually committing to taking every sample would be unwise, as the next sample could make the correct decision obvious, rendering the following $N-1$ superfluous. Thus, we recompute this value at each time step, stopping as soon as

$$\max_N \text{VOI}_t(N) < Nc \tag{14}$$

Derivation of the VOI Finally, we derive a closed-form expression for Equation 12 for arbitrary N. First, note that

$$\text{E}\left[u_{\text{choice}_t}\right] = \max\left\{\mu_t^{(1)}, \mu_t^{(2)}\right\} \tag{15}$$

because one chooses the item with higher expected value, $\mu_t^{(i)}$, and the expected value of that item is simply $\mu_t^{(i)}$. Thus we can rewrite Equation 12 as

$$\text{VOI}_t(1) = \text{E}\left[\max\left\{\mu_{t+1}^{(1)}, \mu_{t+1}^{(2)}\right\}\right] - \max\left\{\mu_t^{(1)}, \mu_t^{(2)}\right\}. \tag{16}$$

Here, $\mu_{t+1}^{(1)}$ and $\mu_{t+1}^{(2)}$ are random variables describing the posterior means of each item after one more sample is collected. $\mu_{t+1}^{(i)}$ is normally distributed with mean $\mu_t^{(i)}$ and standard deviation

$$\zeta_N^{(i)} = \frac{N\tilde{\tau}^{(i)}\bar{\theta}^{(i)}}{\left(\lambda_t + N\tilde{\tau}^{(i)}\bar{\theta}^{(i)}\right)}\sqrt{\frac{1}{\lambda_t} + \frac{1}{N\tilde{\tau}^{(i)}\bar{\theta}^{(i)}}}. \tag{17}$$

where the fraction captures the weight of the sample in the posterior update and the square root term is the standard deviation of the sample given the current posterior mean, taking into account uncertainty in the true value as well as the sampling noise. Finally, using a standard formula for the expectation of the maximum of two Gaussians (Nadarajah & Kotz, 2008), we have

$$\text{E}\left[\max\left\{\mu_{t+N}^{(1)}, \mu_{t+N}^{(2)}\right\}\right] = \mu_t^{(1)}\Phi(\beta) + \mu_t^{(2)}\Phi(-\beta) + \phi(\beta) \tag{18}$$

where

$$\beta = \frac{\mu^{(1)} - \mu^{(2)}}{\sqrt{\left(\zeta_N^{(1)}\right)^2 + \left(\zeta_N^{(2)}\right)^2}}. \tag{19}$$