Evidence or Confidence: What Really Accumulates During a Decision?

Douglas G. Lee[1] Giovanni Pezzulo[2]

[1] California Institute of Technology, Division of the Humanities and Social Sciences

[2] Institute of Cognitive Sciences and Technologies, National Research Council of Italy

Email for correspondence:

DouglasGLee@gmail.com

Abstract

Assessing one's confidence in one's choices is of paramount importance to making adaptive decisions, and it is thus no surprise that humans excel in this ability. However, standard models of decision-making, such as the drift-diffusion model (DDM), treat confidence assessment as a post-hoc or parallel process that does not directly influence the choice – the latter depends only on accumulated evidence. Here, we pursue the alternative hypothesis that what is accumulated during a decision is confidence (that the to-be selected option is the best) rather than raw evidence. Accumulating confidence has the appealing consequence that the decision threshold corresponds to a desired level of confidence for the choice, and that confidence improvements can be traded off against the resources required to secure them. We show that most previous findings on perceptual and value-based decisions traditionally interpreted from an *evidence-accumulation* perspective can be explained more parsimoniously from our novel *confidence-driven* perspective. Furthermore, we show that our novel *confidence-driven DDM* (cDDM) naturally generalizes to any number of decisions – which is notoriously extemporaneous using traditional DDM or related models. Finally, we discuss future empirical evidence that could be useful in adjudicating between these alternatives.

Keywords: choice; confidence; perceptual decision-making; value-based decision-making

Introduction

Humans show an exceptional ability to perform complex behaviors empowered by higher order cognitive processes, such as judgment and goal-directed decision-making. Beyond this great ability, we possess an even more impressive ability to monitor and assess the cognitive processes as they unfold. Known as metacognition, this feature of the human mind allows us to analyze the thoughts that we have and adjust them in a controlled manner so as to tune our own cognitive performance (Fleming et al., 2012). In decision-making, such metacognitive assessment comes in the form of choice confidence, or our subjective belief that what we chose was indeed the correct (or best available) option. Choice confidence is useful in that it can help us learn from our mistakes, or avoid a misallocation of scarce resources (Yeung & Summerfield, 2012). In decisions where there is an objectively correct answer (e.g., perceptual), confidence is highly correlated with accuracy, although it is not always properly aligned (Fleming & Daw, 2017). It has been proposed that different individuals have different levels of metacognitive bias (a systematic mismatch between confidence and accuracy), metacognitive sensitivity (ability to discriminate between correct and incorrect responses), and efficiency (sensitivity conditional on performance; (Fleming & Lau, 2014)). Moreover, in decisions where there is no objectively correct answer (e.g., preferential), people are nevertheless able to report how confident they are about their choices (De Martino et al., 2013). At the neural level, various studies have mapped (retrospective and prospective) aspects of metacognitive ability to (lateral and medial) prefrontal areas and interoceptive cortices (see review in (Fleming & Dolan, 2012)). In spite of the extensive interest and progress in studies of confidence in decision-making, there is as yet no widely accepted computational model that accounts for choice confidence.

As for computation models of decision-making in general, the class of models known as sequential sampling or accumulation-to-bound models is by far the most common. Diffusion decision models (most notably, the drift-diffusion model or DDM) account very well for empirical data across a wide variety of domains, including perceptual and value-based decision-making (Forstmann et al., 2016). In value-based decision-making, the DDM simultaneously explains choice accuracy and response time (RT) distributions dependent on the relative values of the options that compose a choice set (see review in (Ratcliff et al., 2016)). The basic premise of these models is that upon presentation of the choice options, information processing in the decision network of the brain provides a signal about which option is more valuable. This signal is assumed to represent some true value of the options, but because there is noise inherent in neural information processing, the signal is observed over time until the system can reliably declare one option to be more valuable than the other. The core of the DDM is an abstract particle that represents decision evidence (that one option is better than the other) as it accumulates across time. The basic components of the DDM, which relate to the evidence particle, are the *drift rate* at which it moves on average (proportional to the true value difference of the options), the *diffusion* parameter by which it is momentarily perturbed (the extent of signal corruption in the system), and the *evidence threshold* which halts the particle as a choice is made (the minimum level of evidence required to declare that one option is better than the other). It has been proven that the DDM can provide the optimal solution to simple decision-making problems, where prior beliefs about the option values are updated in a Bayesian manner through sequential evidence sampling, until a choice can confidently be made based on posterior value beliefs (Fudenberg et al., 2018; Tajima et al., 2016). That work shows that the decision thresholds

should diminish over deliberation time[1], in any situation where the difficulty of the task is not known *a priori* and there is some cost associated with evidence accumulation (as would be the case in most preferential choice contexts). As the accumulator is here defined as the *difference* in the posterior beliefs about the option values, the height of the threshold at the time of choice could directly provide a measure of *choice confidence* (i.e., the probability that the chosen option was better (Pouget et al., 2016)). In this case, the collapse of the threshold as time lingers could be interpreted as a tradeoff between accepting a lower-than-desired level of confidence about the choice in exchange for a conservation of deliberation resources (e.g., time, metabolites, neural capacity).

One problem with using the threshold height as a direct readout of choice confidence is that it does not take into account the *precision* of the posterior value estimates. In the standard DDM, value estimates are point estimates, and do not consider the uncertainties of these estimates. A standard measure of uncertainty for (Gaussian) probability distributions is the variance of the distribution, or its inverse precision. Specifically, with flat priors and Bayesian sequential updating, the difference in the posterior means of a pair of options could be equivalent to that of another pair, but with different levels of precision. For example, the posterior estimates of one option pair could have been formed from relatively inconsistent evidence samples (i.e., many samples in support of either option, hence lower-precision posteriors), whereas the posterior estimates of the other option pair could have been formed from relatively consistent evidence samples (i.e., most samples in support of the same option, hence higher-precision posteriors; see Figure 1). In this case, both choices would reach the same evidence threshold, but

[1] The collapsing threshold is a hallmark of the optimal rendition of the DDM, though it is not a feature of the standard DDM. However, some non-optimal versions of the DDM have also included collapsing thresholds (Voskuilen et al., 2016).

it seems intuitive that the option pair with the higher posterior precision should correspond to higher choice confidence. In other words, the idea of using a threshold as a measure of confidence does not fully account for a fundamental psychological dimension of confidence — the fact that people are more confident when they have more precise information. To rectify this, the threshold height at choice would need to be transformed (e.g., scaled by precision and passed through a sigmoidal function) in order to calculate confidence. However, although this might solve the problem mathematically, it would seem an *ad hoc* solution rather than a principled one, and it would then remain unclear as to why the deliberation process would seek a target (threshold) that would sometimes register as high confidence and sometimes low (setting aside the issue of the eventual threshold collapse). It seems more reasonable to believe that the same (high) level of confidence would always be the default target for similar decisions.

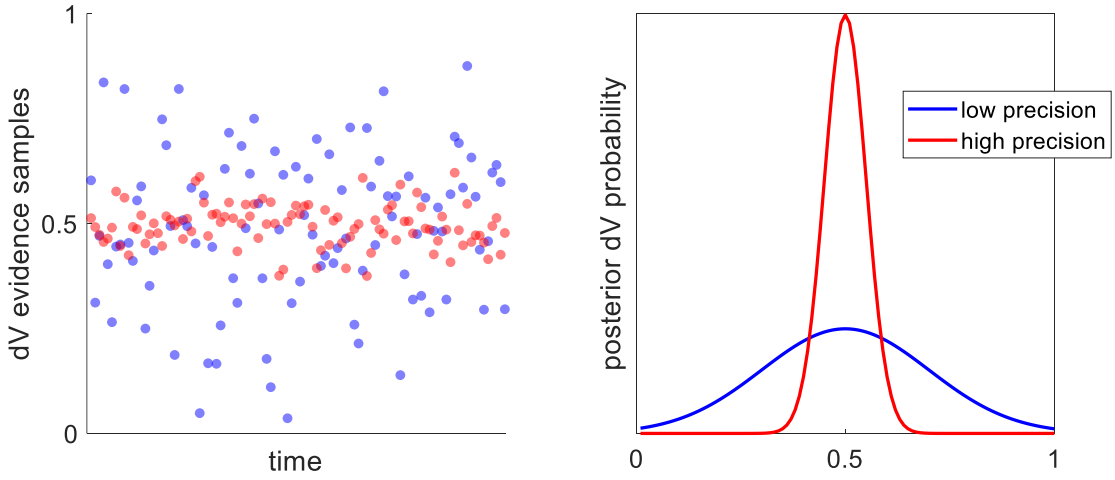

Figure 1: *Posterior estimates of the difference in option values (dV) might have similar means but very different levels of precision. The left panel shows sample streams of evidence for two different choice trials (one in blue, one in red). Notice that the samples fluctuate around the same average value, but with more (blue) or less (red) variability. The right panel shows what the posterior dV estimates for the two trials would look like—both have the same mean, but the*

red estimate has a much higher precision than the blue estimate. This higher-precision estimate should instill greater choice confidence.

Recent work has attempted to use the DDM to predict choice confidence in addition to choice and RT (Calder-Travis et al., 2020; Drugowitsch et al., 2019; Kvam & Pleskac, 2016; Moran et al., 2015; Moreno-Bote, 2010; Pleskac & Busemeyer, 2010; Yeung & Summerfield, 2012). However, most such attempts have required additional assumptions and components on top of the DDM itself, which detracts from the elegance of the original model and its ability to so reliably account for key decision variables. Given the success of the DDM in accounting for other key aspects of choice (i.e., the choice itself and the RT), and assuming that confidence about a choice is inherently linked to those other aspects (Kiani & Shadlen, 2009; Van Den Berg et al., 2016), one might conclude that the DDM should be able to simultaneously account for all three variables without the need for any ad-hoc features. Calder-Travis et al (2016) conducted a thorough analysis of various sorts of DDM found in the literature, specifically focused on their ability to predict choice confidence in a variety of different datasets. However, those authors chose not to perform their model comparison while simultaneously accounting for choice, RT, and confidence. Furthermore, they only examined data from perceptual decision studies, so it is unclear if their findings generalize to preferential decisions.

Here, we pursue an alternative hypothesis that the response threshold in the DDM should directly represent a target level of choice confidence, and thus that the accumulator particle should represent the momentary level of confidence across time. All this would require is that the particle be calculated not simply as the difference in posterior value estimates, but rather a sigmoidal transformation of that difference scaled by the posterior precision (i.e., the normative

posterior probability that one option is better than the other, given the accumulated evidence thus far). Hence, the version of DDM that we propose would pursue a strategy opposite to that of previous models: whereas previous models have transformed the decision variable after the choice was already made in order to calculate confidence, we propose that the transformation from evidence to confidence occurs during (and is indeed an integral component of) the accumulation process. Furthermore, this allows for the posterior precision to be incorporated dynamically throughout deliberation, which would preclude the shortcomings described above.

Another shortcoming of existing sequential sampling models is their inability to account for the impact of value estimate uncertainty. Decision makers form not only estimates of value for the different options that they consider, but also judgments of (un)certainty about those value estimates (De Martino et al., 2013; Gwinn & Krajbich, 2020; Lee & Daunizeau, 2020; Lee & Coricelli, 2020; Lee & Daunizeau, 2021; Polanía et al., 2019). These feelings of certainty are important, because they alter choice behavior (Lee & Daunizeau, 2020; Lee & Coricelli, 2020; Lee & Daunizeau, 2021). Most existing versions of the DDM (or any other sequential sampling model) exclude the possibility that option-specific value certainty might play a role in the decision process. However, it has recently been shown that the DDM indeed provides a better explanation of choice data when the drift rate is adjusted to reflect the different degree of certainty that the value of each option has (Lee & Usher, 2020). The authors demonstrated that the so-called *signal-to-noise DDM* (snDDM) is capable of accounting for the positive impact of option-specific value certainty on choice accuracy and confidence, and the negative impact on RT (Lee & Coricelli, 2020; Lee & Daunizeau, 2021).

Other recent work showed that the DDM fits better to empirical data when the drift rate is modeled as a sigmoidal (rather than linear) function of the value difference between the options

(Fontanesi et al., 2019). While the authors proposed the non-linear mapping to account for inter-individual differences in sensitivity to value difference, we believe it is equally apt for describing inter-trial differences within the same decision maker. The logic can easily be understood if one considers the momentary relative evidence signal, which is meant to provide information about which option is better, but not explicitly about the magnitude of the values (or the value difference). Therefore, an increment of evidence that very highly favors one option might not provide much more information than an increment of evidence that moderately favors that same option. If we allow the DDM drift rate to be modulated by value certainty (as in the snDDM) as well as transformed by a sigmoid function, the result is that the momentary signal itself traces the evolution of choice confidence across deliberation. The threshold that terminates the process thus becomes the target level of confidence that the decision maker expects to achieve before choosing an option (see Lee & Daunizeau, 2021).

We propose that the signal that is monitored during a diffusion decision process is not merely a cumulative tally of relative evidence for the options (that is later transformed into a belief about confidence), but explicitly the momentary confidence about knowing which option is best. A decision can be then terminated when a target confidence is reached. This is different from the classic DDM, where value signals about the options drive the evidence accumulator until it reaches some arbitrary threshold, at which point a Bayesian calculation of confidence might transform the relative evidence signal into a choice confidence signal via a sigmoid function of the evidence scaled by the certainty about the evidence. With our proposed *confidence-driven DDM* (cDDM), the signal that is monitored has already undergone the requisite transformation. Such a proposition would be parsimonious for a variety of reasons. For one, the cDDM could be more easily adapted to decisions between any number of options, as the

momentary confidence signal would not be confined to a comparison between two values (see Appendix); standard DDMs only apply to binary decisions (but see Kvam, 2019). Furthermore, the cDDM could be directly applied to potentially any type of decision without having to adjust the meaning of its components or substantially alter its parameters (e.g., perhaps the target confidence level for a decision maker might be similar across decision domains, whereas the relative evidence requirements might grossly differ). Finally, from a computational viewpoint, it would seem more useful (or parsimonious) for the brain to assess the reliability of the relevant information (i.e., confidence about which option seems better) as a choice response freely develops, rather than simply keeping a tally of how many stochastic samples favor each option, stopping when some arbitrary threshold has been reached, and only then checking to see how confident it is about the choice that would have already been made. Our proposal is in line with previous work showing that decision makers often change their minds about which option is better during the course of deliberation, and that such changes of mind coincide with changes of confidence (Van Den Berg et al., 2016). This suggests that confidence is indeed continuously monitored during deliberation, with changes of mind being triggered whenever confidence (that the current to-be chosen option is the best) disappears.

In addition, the cDDM includes a "collapsing bound" mechanism, which encourages the decision process to terminate even if the target level of confidence was not yet achieved - and hence to choose an option with a lower level of confidence, because further effort (cognitive, metabolic, or otherwise) to increase confidence is deemed too costly to continue (see Lee & Daunizeau, 2021). This is analogous to what happens in the optimal DDM, where the threshold collapses explicitly due to a cost of deliberation that accumulates over time (Fudenberg et al.,

2018; Tajima et al., 2016; see Malhotra et al., 2018 for an alternative rationale for collapsing thresholds)[2].

Thus, the decision process of the cDDM can be summarized as a continuous readout of a value signal for each choice option, the stability of which drives a continuous readout of choice confidence towards a predetermined target level, which is gradually reduced in consideration of the growing resource expenditure or urgency of the decision (see Figure 2 for an illustrative example).

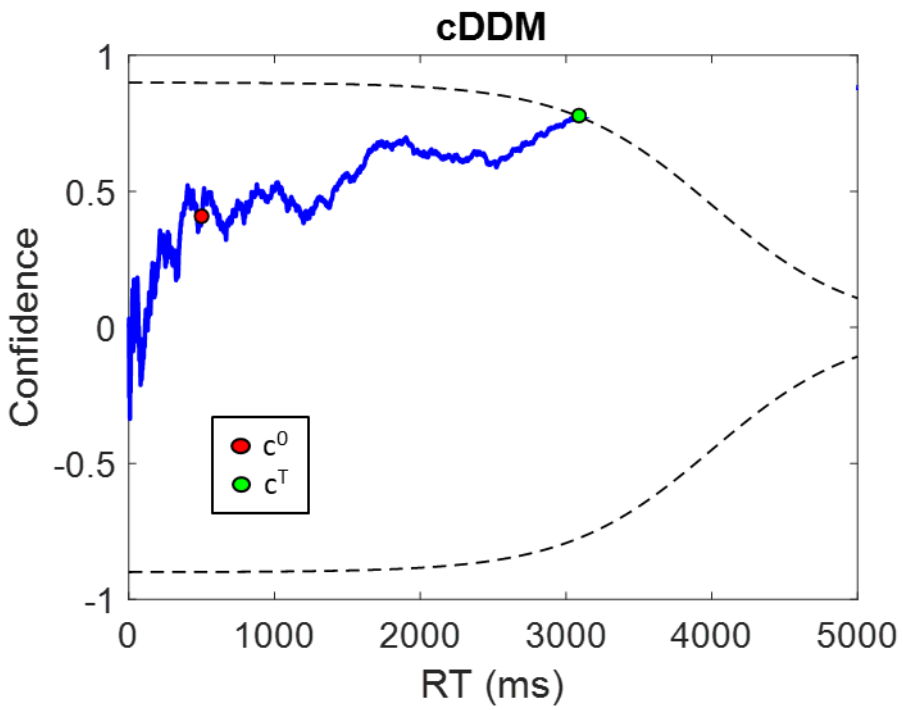

Figure 2: *An illustrative example of the cDDM process. The blue trace illustrates the evolution of choice confidence across response time (RT). At t=0, the options are unknown, and thus confidence = 0. For the first 500 ms (shown to be longer than the typical 200 ms for visualization purposes), the options are being perceived and identified, while imprecise value*

[2] An alternative (but mathematically equivalent) perspective is that the decision mechanism includes an *urgency signal* that amplifies the incoming decision evidence, which (akin to the collapsing bound mechanism) encourages an end to deliberation and a final choice response without excessive further delay (Churchland et al., 2008; Thura & Cisek, 2014; van Maanen et al., 2016).

representations form in parallel. This leads to an initial confidence level of c^0 by the time deliberation begins. Momentary precision-weighted value evidence drives confidence until it reaches a threshold and a choice is made (here, for option 1). Note that the target confidence threshold collapses over time, as the agent recognizes the importance of conserving resources (e.g., time and cognitive effort) rather than continuing to increase confidence.

MODEL

In this paper, we propose a novel version of a drift-diffusion model that involves only minor alterations to the core mathematics, but major alterations to the core assumptions and interpretations. Taken together, these alterations allow our new cDDM to directly account for choice confidence alongside choice accuracy and response time, in an integrated manner. Below, we first summarize the standard mathematical formulation of the standard DDM, complete with some well-known optional parameters (starting point bias, non-decision time, collapsing bounds). We then describe the alterations that distinguish the cDDM from the traditional DDM, as well as their contrasting assumptions and interpretations.

In the standard DDM, the decision deliberation process begins with both time and cumulative evidence equal to zero. At each time step (t) after the onset of deliberation, evidence (e) for each choice alternative (i) is drawn from an independent Gaussian distribution whose mean (μ_i) is the true value of the option and whose variance (σ^2) represents the amount of signal corruption due to the inherent stochasticity of neural processing (which is the same for all options):

$$e_{i,t} \sim N(\mu_i, \sigma^2)\ , \forall\ i \in \{1,2\}$$

The evidence for each option at each time step is compared and the balance is added to the cumulative total evidence (x) for one option over the other:

$$x_t = x_{t-1} + e_{1,t} - e_{2,t}$$

$$x_0 = 0$$

The deliberation process proceeds until the accumulated evidence reaches one of two symmetric predetermined threshold bounds (b>0 for option 1, -b for option 2). It is possible to bias the starting point of the process (x_0) away from zero, but this is arbitrary unless there is a good reason to assume that one option is *a priori* more likely to be better (e.g., instruction about or observation of an asymmetrical environment during a sequential decision-making task). It is also possible to include a measure of non-decision time (NDT) to represent the non-zero time that it takes for the brain to process perceptual information while recognizing the options or motor control information while entering a response, which is simply added to deliberation time (DT) to calculate total response time (RT). As mentioned above, one possible way to map from evidence to confidence in the traditional DDM would be to pass it through a sigmoidal transformation. Figure 3 illustrates simulated evidence accumulation, and the corresponding confidence transformation, for 100 trials of varying difficulty (i.e., unsigned value difference between the options).

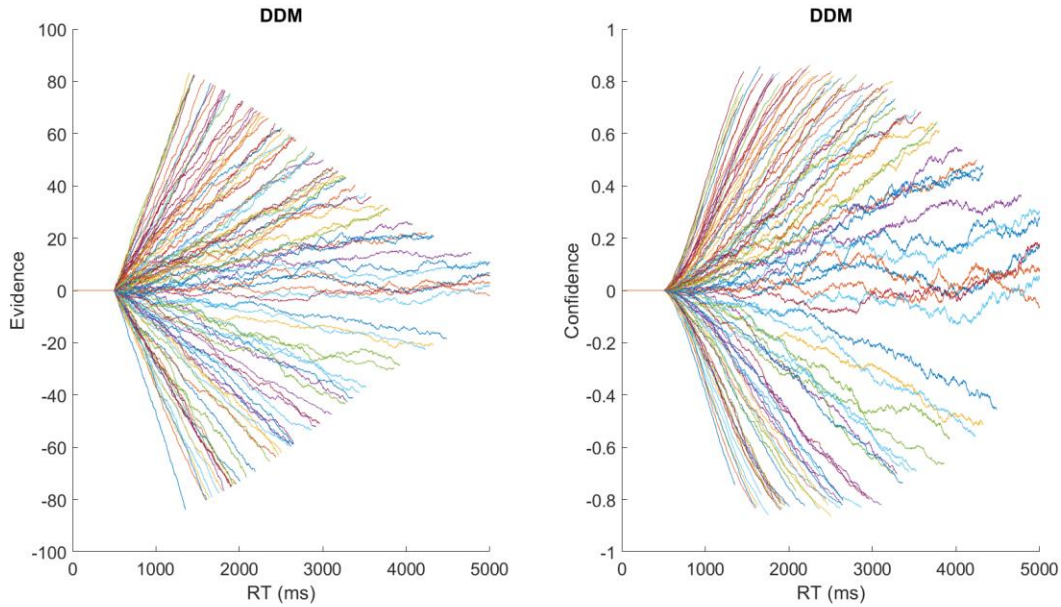

Figure 3: *An illustrative example of 100 simulated trials of varying value difference (dV), based on a standard DDM with collapsing bounds. Each colored curve represents one trial. The left plot shows the accumulation of evidence across time, where each process terminates upon reaching one of the bounds. The right plot shows how the evolution of confidence would look if it were a moment-by-moment sigmoidal transformation of the cumulative evidence.*

In the cDDM, the momentary variable that is monitored is not merely a relative value signal for the options (as in the standard DDM), but rather an assessment of confidence (c_t) that considers both the value estimates of each option (μ_i) and the inverse precision of those value estimates (σ_i^2), passed through a sigmoidal transformation to compute the probability that one option is better than the other:

$$e_{i,t} \sim N\left(\frac{\mu_i}{\sigma_i^2}, \sigma^2\right), \forall\, i \in \{1,2\}$$

$$v_{i,t} = \sum_{T=1}^{t} \frac{e_{i,T}}{t}$$

$$v_{i,0} = 0$$

$$p_t = \frac{t}{\sigma^2}$$

$$c_t = \frac{2}{1 + e^{\beta(v_{1,t} - v_{2,t})\sqrt{p_t}}} - 1$$

where $v_{i,t}$ is the momentary value assessment of option i at time t and p_t is the momentary precision of the value comparison. Here, the standard sigmoid function undergoes an affine transformation to ensure that the incremental evidence falls within the range (-1,1), with 1 indicating maximal confidence that option 1 is correct and -1 indicating maximal confidence that option 2 is correct, and β is an agent-specific free parameter to control the rate at which evidence augments confidence.

The cDDM assumes that during the pre-deliberation NDT, when the options are being perceived and identified, a low-precision value representation of the options is automatically formed (cf. (Lebreton et al., 2009, 2015)). Here, NDT corresponds to the time required for the perception and identification of the options (we choose not to consider post-deliberation motor response NDT). Crucially, the cDDM assumes that during this period, value representations of the options are formed in parallel, but with rather low precision. This is because in this early sensory processing stage, the options are not yet completely perceived or identified. This early representation formation can be formalized in the same manner as during explicit deliberation, only with lower precision ($\sigma_{NDT}^2 > \sigma^2$; e.g., lower attention will be assigned to the valuation task at this stage):

$$e_{i,t} \sim N\left(\frac{\mu_i}{\sigma_i^2}, \sigma_{NDT}^2\right), \forall\, i \in \{1,2\}, 0 < t \leq NDT$$

At the end of NDT, the agent will have an initial estimate of the value (v) of each option, as well as an initial estimate of precision (p) for the value estimates:

$$v_i^0 = \sum_{t=1}^{NDT} \frac{e_{i,t}}{NDT}, \forall\, i \in \{1,2\}$$

$$p^0 = \frac{NDT}{\sigma_{NDT}^2}$$

Based on the initial estimate and precision of value for each option, the agent will already have an initial sense of confidence by the time deliberation commences. Thus, whereas the confidence monitor (c) begins the decision process at zero, it typically arrives at a non-zero point by the end of the NDT. This creates a starting point bias (z) equal to the subjective probability that one of the options is better than the other (i.e., confidence) prior to explicit deliberation:

$$z = c^0 = \frac{2}{1 + e^{\beta(v_1^0 - v_2^0)\sqrt{2p^0}}} - 1$$

More generally, both the momentary value estimates and the momentary confidence level will continue to evolve across deliberation time. These variables will incorporate both the information that was automatically processed during NDT and the information that was deliberately processed after NDT:

$$v_i^T = \sum_{t=1}^{T} \frac{e_{i,t}\pi_t}{p^T}, \forall\, i \in \{1,2\}$$

$$\pi_t = \begin{cases} \sigma_{NDT}^{-2} \; if \; 0 < t \leq NDT \\ \sigma^{-2} \; if \; t > NDT \end{cases}$$

$$p^T = \sum_{t=1}^{T} \pi_t = \frac{NDT}{\sigma_{NDT}^2} + \frac{T - NDT}{\sigma^2}$$

$$c^T = \frac{2}{1 + e^{\beta(v_1^T - v_2^T)\sqrt{p^T}}} - 1$$

To describe the cDDM as a Markov decision process, the equations for the evolution of value (v) and its precision (p) can be written as the following difference equations:

$$v_{i,t} = \frac{v_{i,t-1}p_{t-1} + e_{i,t}\pi_t}{p_t}\ , \forall\ i \in \{1,2\}\ , t > 0$$

$$p_t = p_{t-1} + \pi_t$$

$$c_t = \frac{2}{1 + e^{\beta(v_{1,t} - v_{2,t})\sqrt{p_t}}} - 1$$

The decision process will continue to step forward in time and acquire additional samples of value evidence until confidence reaches a predetermined boundary (b or –b). If the process terminates at c=b, option 1 will be chosen with level of confidence b; if it terminates at c=-b, option 2 will be chosen with confidence b. The final inferred values of the options will be v_i^T, where T is the time at which the process terminated. Figure 4 illustrates simulated value estimate refinement, and the corresponding confidence accumulation, for 100 trials of varying difficulty.

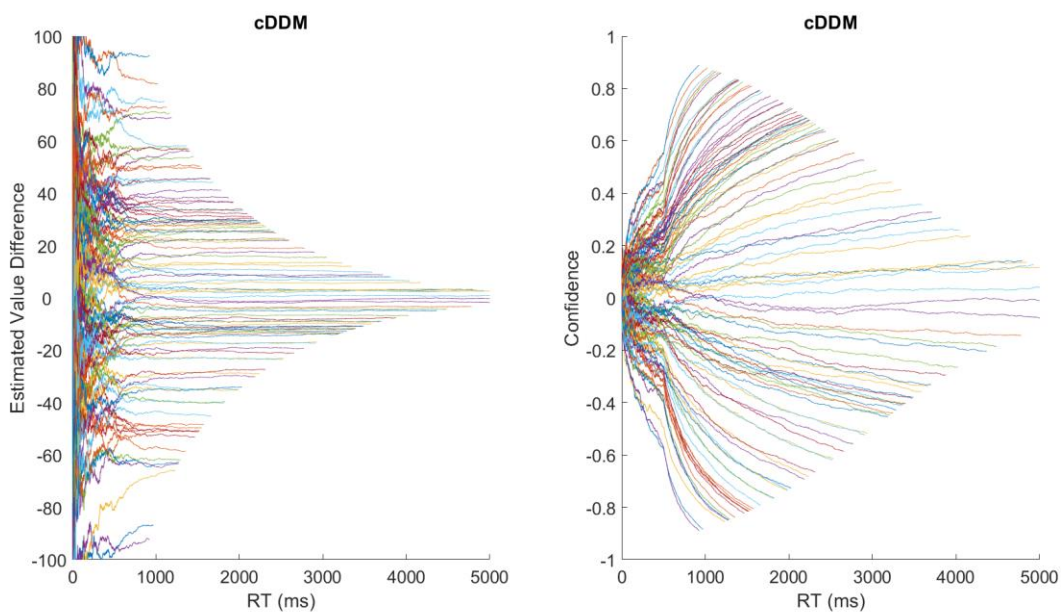

Figure 4: *An illustrative example of 100 simulated trials of varying value difference (dV), based on the cDDM. Each colored curve represents one trial. The right plot shows the accumulation of confidence across time, where each process terminates upon reaching one of the bounds. The left*

plot shows how the evolution of estimated value difference would look if it were to be extracted from the confidence signal.

DISCUSSION

It is widely assumed that decision-making consists of an accumulation-to-bound process, during which decision evidence (that one option is better than the other) is accumulated over time until a threshold is met (Gold & Shadlen, 2007). This hypothesis is implicit in many mainstream models of decision-making and in particular the DDM, which has been highly successful in explaining choice and reaction time data across a very large number of studies (Ratcliff et al., 2016), but which requires additional assumptions to be able to simultaneously explain choice confidence. Here, we advanced the alternative hypothesis that what is accumulated during a decision is the momentary level of confidence across time, and that the decision threshold thus corresponds to a target level of choice confidence. We incorporated this assumption in a novel model (the cDDM) that follows a similar accumulation-to-bound approach as does the DDM, but which monitors a confidence signal (i.e., the posterior probability that one option is better than the other, given the accumulated evidence thus far) rather than the accumulated evidence *per se*. This proposal is related to that of Lee and Daunizeau (2021), who argued that deliberation itself is a form of effort-confidence tradeoff as described above (regarding collapsing decision thresholds), but described it from the other side of the equilibrium point—a decision maker will choose to invest effort in expectation of achieving higher confidence (compared to choosing prematurely; see also Chaiken et al., 1989). Thus, during deliberation, incremental mental effort drives the drifting confidence particle until it reaches a threshold and the process terminates with a choice. Another related idea has been proposed in the realm of risky choice (Navarro-Martinez et al., 2018). Here, evidence samples based on probabilistic

utility functions serve as input into a t-test calculation (mean / standard error of the mean) and the decision maker stops deliberating when the t-score reaches a target confidence level.

We argue that the novel "confidence accumulation" proposal has computational advantages over the standard "evidence accumulation" point of view, and that it permits us to recast several existing empirical findings in novel terms. Furthermore, our approach allows us to make some novel empirical predictions, which we discuss below.

Computational advantages of cDDM

From a computational perspective, one may ask, "What is more worthwhile to accumulate in diffusion decision models: evidence or confidence"? Accumulating confidence rather than evidence entails three main computational advantages.

First, the cDDM includes a built-in mechanism that permits accounting for not only the evolving balance of evidence about the values of the different options, but also a measure of certainty about those values. This allows for a more direct readout of confidence at the time a choice is made, relative to the standard DDM (which ignores the concept of value certainty), and to take into account the *precision* of the posterior value estimates. Moreover, it also allows for a direct readout of confidence at any time prior to the choice (during deliberation). Thus, the decision apparatus can distinguish from situations where the instantaneous balance of evidence is equal but the precision of that evidence differs, which would allow for the system to optimally determine whether it should continue to deliberate or not. Specifically, in situations where the accumulating evidence has lower precision, deliberation should persist for longer (cf. Lee & Usher, 2020).

Second, a decision mechanism that accumulates confidence (like the cDDM) generalizes more easily to decisions between any number of options, compared to one that accumulates

evidence (like the standard DDM). The DDM was originally conceived to model decisions between two alternatives. While it could be extended to the account for decisions between multiple alternatives, such an extension would require introducing multiple decision variables, which might become implausible as the number of choice options increases (Churchland et al., 2008; Roxin, 2019). On the contrary, the cDDM can model multi-alternative choices using a single decision variable that tracks the difference between the posterior probabilities of the best and the next-best options (see Appendix; see Li & Ma, 2020 for an alternative approach). As such, the cDDM can handle arbitrarily-large choice sets without adding much complexity to the computations involved.

Finally, the cDDM rationally trades off confidence (the target) and effort (to increase confidence towards the target) during the choice. For example, in conditions where an excessive amount of time is required to increase one's confidence or when time is lapsing without any increase in confidence, it may be optimal to decide immediately (even randomly) rather than spend additional time deliberating. To afford this cost-benefit tradeoff, the brain needs to monitor another variable throughout the deliberation process: a resource expenditure monitor[3] that represents the total amount of cognitive resources that have been invested in the current decision across time. As resource expenditure approaches some critical level, where the anticipated benefit of further expenditure diminishes towards the anticipated cost, this signal would encourage the decision proces(Shenhav et al., 2017; Zenon et al., 2019)s to terminate even if the target level of confidence were not yet achieved. In the optimal DDM and the cDDM, this can be achieved by introducing a threshold that collapses due to a cost of deliberation that

[3] There is debate about what resources are consumed in a costly manner during cognitive processes, including metabolites, time, or capacity (Zenon et al., 2019; see Shenhav et al., 2017 for a review). For our purposes, we remain ambivalent and use the term resources in a more abstract sense (i.e., whatever the true costly resource turns out to be, our model will not change).

accumulates over time (Fudenberg et al., 2018; Tajima et al., 2016). An alternative suggestion is that the brain uses an multiplicative urgency signal[4] that encourages an end to deliberation and a final choice response without excessive further delay (Churchland et al., 2008; Thura & Cisek, 2014; van Maanen et al., 2016). Computational models have been proposed in which the evidence accumulator is amplified at a rate that increases across time, representing the urgency signal (Cisek et al., 2009; Ditterich, 2006). In the DDM, an amplified drift rate is mathematically equivalent to a diminished threshold (if the noise term is adjusted accordingly), so the purported urgency signal might be functionally identical to the collapsing bound in the optimal DDM. In either case, it is important to consider that choice in the cDDM (or in DDM models that use collapsing bounds, or in urgency-gating models) is not a simple function of evidence but also of the costs (e.g., cognitive resources or time) of deliberation.

Reinterpreting Empirical Findings through the Lens of the cDDM

Whether confidence is monitored during (or in parallel with) a choice or after it has been the object of a longstanding debate (Fleming & Daw, 2017). However, many recent studies suggest that confidence judgments influence the development of choice responses (Schulz et al., 2021) rather than just being post-hoc readouts. For example, people who already feel confident about their choice after accumulating a certain amount of evidence do not accumulate additional evidence (or they process only choice-congruent information from that point forward; Rollwage et al., 2020). Furthermore, different people have different rates of urgency during a choice (which can equivalently be interpreted as having different rates of threshold collapse, meaning

[4] Urgency signals are typically presented with respect to time specifically, where responses are made sooner in order to save time. Rising deliberation costs and urgency are not mutually exclusive, and in fact the latter can be considered a special case of the former. Furthermore, one could monitor both time urgency and other costs of deliberation, and diminish the confidence threshold in response to either or both.

that they are more or less willing to accept lower confidence in exchange for saving effort; Hauser et al., 2017). All these studies (and others) suggest that the monitoring of confidence is part and parcel of the decision process and influences it in real time. A possible way to conceptualize these findings is that the decision process consists of two parallel processes: an evidence accumulation process that is responsible for the choice and a confidence monitoring process (carried out with a slight delay, since it needs to take as input the momentary state of evidence accumulation) that can influence the evidence accumulation (e.g., by setting collapsing bounds). Our proposed cDDM model goes even further and proposes that decision-making only requires a single (confidence accumulation) process.

Beyond real-time confidence monitoring, there is already some strong evidence that people explicitly use confidence to tune latent parameters such as those in the DDM. In a recent study, participants were asked to make their choices (distinguish between Gabor patches of different orientations) such that they would expect to achieve 70%, 85%, or 90% accuracy across trials (Balsdon et al., 2020). Participants were indeed able to tune their behavior in this way, decreasing RT for lower target confidence and increasing RT for higher target confidence, with subjective confidence levels being in line with objective accuracy percentages. Furthermore, when allowed to respond without any externally-imposed confidence target, participants generally stopped accumulating evidence much earlier and achieved accuracy/confidence levels that were lower than what they were demonstrably capable of (in the high target confidence condition). This supports our claim that decision makers set a target confidence level and stop accumulating evidence once that confidence level has been reached. Additional empirical evidence in support of this was previously shown in the realm of probabilistic decision-making, where participants paid a cost to sequentially reveal cues but stopped whenever the probability of

knowing the correct answer rose above a (participant-specific) threshold (Hausmann-Thürig & Läge, 2008). Notably, the same behavior held even in a condition where participants knew that additional cues could not change the relative rankings of the response options, suggesting that increasing confidence was itself an independent goal.

At the neurobiological level, neural signals reflecting (or necessary for) confidence accumulation have been identified during perceptual decisions in parietal (LIP) areas in a study by Kiani and Shadlen (2009). The fact that this representation is found during the decision period and not (only) afterwards suggests that confidence is not a post-hoc computation but can be integral to the choice. The task used by these authors consisted of a standard ("left" or "right") decision; but to test the effects of confidence on the choice, it also included a "post-decision wagering" condition in which monkeys could select a lower-reward but "sure" offer. The results show that monkeys indeed choose the sure offer when they are (or should be) less confident about the main ("left" or "right") choice. While Kiani and Shadlen appeal to a standard evidence-accumulation model to explain LIP firings, what they actually use to fit the behavioral results is an "extension" of evidence accumulation: a model where the standard decision variable is replaced by the log-odds of a correct decision, which (similar to the cDDM) is a representation of choice confidence (Kiani & Shadlen, 2009). This leaves us with the slightly uncomfortable situation where the LIP area is responsible for the "left" versus "right" choice (which is empowered by accumulating evidence) but some downstream process is necessary to calculate the log-odds to make the "sure" choice. A more parsimonious explanation would describe the entire choice (between the three options) as a confidence accumulation process, rather than as an accumulation of evidence followed by a downstream computation of confidence.

A similar argument can be made in relation to another study that perturbed decision states in real time, and which predicted the decision variable using a confidence (log-odds) model for one versus the other choice option (Peixoto et al., 2021). Similar to the previous study, the experimental results are explained in a more intuitive way by appealing to a model that monitors confidence compared to a model that encodes evidence (with confidence resulting only as a byproduct).

Along similar lines, it is possible to reformulate several previous findings that were initially interpreted as evidence for the DDM in terms of the cDDM. For example, a study of multi-alternative choice reveals that the average LIP neuron firing rate after target presentation but before motion onset is lower for choices between four alternatives compared to those between two alternatives (Churchland et al., 2008). This is consistent with a neural signal for confidence, since the *a priori* probability of the correct response is lower (25% versus 50%).

Finally, it is worth noting that the brain may monitor confidence across multiple trials, not just within trials. A study reports that confidence information can be carried forward across multiple-step decisions, “bridging” the individual steps (i.e., the intermediary decisions) together (van den Berg et al., 2016). A neural study reports that the belief that stimuli are biased (here, in terms of the net direction of dot motion) evolves within blocks, plausibly as a function of confidence about decisions in individual trials (Zylberberg et al., 2018). Interestingly, the form of this confidence update within blocks (see their Figure 2D) is the same as the form of the confidence update that occurs in the cDDM within trials (see our Figure 4B). This equivalence can be useful methodologically, as one could replace confidence updates within trials (which are difficult to estimate) with confidence updates across trials (which are easier to estimate).

Novel empirical predictions of the cDDM

The results of the aforementioned studies are compatible with the cDDM, but do not unambiguously distinguish it from alternative models. Hence, fully assessing the cDDM in comparison to alternative theories requires testing situations that are more diagnostic of one theory or the other. A bonus feature of the cDDM, which is not possible with previous versions of DDM, is that it allows for the exposition of *choice-induced confidence change* (a measure not previously reported in the literature, but see (Fleming & Daw, 2017)). We introduce this new terminology as a label for the phenomenon that decision makers are generally more confident about their choices when given ample amount of time to consider the options before responding, compared to situations where they are required to respond immediately. This would capture the gain (or sometimes, loss) of confidence that the agent obtains via deliberation (Δc):

$$\Delta c = \begin{cases} c^T - c^0 \; if \; c^0 > 0 \, , c^T > 0 \\ -c^0 - c^T \; if \; c^0 > 0 \, , c^T < 0 \\ c^0 - c^T \; if \; c^0 < 0 \, , c^T < 0 \\ c^0 + c^T \; if \; c^0 < 0 \, , c^T > 0 \end{cases}$$

$$= |c^T| - |c^0|$$

One previous study has reported a finding that could potentially also merit the label choice-induced confidence change (Wokke et al., 2020). That study showed that people reported higher levels of confidence after having committed to their choice responses, relative to a condition where they first reported confidence about knowing the correct choice before actually reporting their choice. However, this suggests a sort of post-choice confidence bias akin to the theory of cognitive dissonance resolution (Festinger, 1957), rather than an instrumental augmentation of confidence that occurs during deliberation. Future studies would be required to more fully examine the two possible causes of choice-induced confidence change. While the mechanisms of choice-induced confidence change remain to be tested in future studies, we note

that this is exactly the mechanism that may be needed to form beliefs about stimulus direction within blocks, as a function of confidence about decisions in individual trials (Zylberberg et al., 2018).

Appendix

Define choice confidence P_c as the probability that the (predicted) experienced value of the (to-be) chosen option is higher than that of all of the (to-be) unchosen options. Critically:

$$P_c = \begin{cases} P\left(V_1 > V_2 \,\&\, V_1 > V_3\right) & \text{if item \#1 is chosen} \\ P\left(V_2 > V_1 \,\&\, V_2 > V_3\right) & \text{if item \#2 is chosen} \\ P\left(V_3 > V_1 \,\&\, V_3 > V_2\right) & \text{if item \#3 is chosen} \end{cases}$$
$$= \begin{cases} P\left(V_1 > \max[V_2, V_3]\right) & \text{if } \mu_1 > \max[\mu_2, \mu_3] \\ P\left(V_2 > \max[V_2, V_3]\right) & \text{if } \mu_2 > \max[\mu_1, \mu_3] \\ P\left(V_3 > \max[V_1, V_2]\right) & \text{if } \mu_3 > \max[\mu_1, \mu_2] \end{cases}$$

where the second set of equations derives from assuming that the option with the highest expected value will be chosen. Choice confidence is thus derived from a comparison between the best option (i.e., the option with the highest expected value μ_{best}) and the maximum of the rest of the options. For example, for a choice set containing three options, with the first option having the highest expected value, the expected maximum value of the other options is (Nadarajah & Kotz, 2008):

$$V_{max} = max(V_2, V_3) \sim N(\mu_{max}, \sigma_{max})$$

$$\mu_{max} = E[max(V_2, V_3)] = \int x f(x) \partial x$$

$$f(x)\text{: } pdf\ of\ V_{max} = pdf\left(\frac{x-\mu_2}{\sigma_2}\right) * cdf\left(\frac{x-\mu_3}{\sigma_3}\right) + pdf\left(\frac{x-\mu_3}{\sigma_3}\right) * cdf\left(\frac{x-\mu_2}{\sigma_2}\right)$$

For a choice set with four options, with the first option having the highest expected value, an iteration of this procedure will provide the expected value of the maximum of the options other than the best:

$$\mu_{max} = E[max(max(V_2, V_3), V_4)] = E\left[max\left(V_{max_{2,3}}, V_4\right)\right]$$

The first and second moments of the maximum function can be calculated iteratively to obtain first the expected maximum value of two of the three non-best options, then the expected maximum value of this value and the third non-best option:

$$\mu_{max_{2,3}} = E\left[V_{max_{2,3}}\right] = \mu_2 * \ cdf\left(\frac{\mu_2 - \mu_3}{\theta}\right) + \mu_3 * \ cdf\left(\frac{\mu_3 - \mu_2}{\theta}\right) + \theta * pdf\left(\frac{\mu_2 - \mu_3}{\theta}\right)$$

$$\theta = \sqrt{{\sigma_2}^2 + {\sigma_3}^2}$$

$$E\left[{V_{max_{2,3}}}^2\right] =$$

$$({\sigma_2}^2 + {\mu_2}^2) * cdf\left(\frac{\mu_2 - \mu_3}{\theta}\right) + ({\sigma_3}^2 + {\mu_3}^2) * cdf\left(\frac{\mu_3 - \mu_2}{\theta}\right) + (\mu_2 + \mu_3) * \theta * pdf\left(\frac{\mu_2 - \mu_3}{\theta}\right)$$

$${\sigma^2}_{max_{2,3}} = E\left[{V_{max_{2,3}}}^2\right] - {\mu_{max_{2,3}}}^2$$

$$\downarrow$$

$$E[V_{max}] = \mu_{max_{2,3}} * \ cdf\left(\frac{\mu_{max_{2,3}} - \mu_4}{\theta}\right) + \mu_4 * \ cdf\left(\frac{\mu_4 - \mu_{max_{2,3}}}{\theta}\right) + \theta * pdf\left(\frac{\mu_{max_{2,3}} - \mu_4}{\theta}\right)$$

$$\theta = \sqrt{{\sigma_{max_{2,3}}}^2 + {\sigma_4}^2}$$

$$E\left[{V_{max}}^2\right] = \left({\sigma_{max_{2,3}}}^2 + {\mu_{max_{2,3}}}^2\right) * cdf\left(\frac{\mu_{max_{2,3}} - \mu_4}{\theta}\right) + ({\sigma_4}^2 + {\mu_4}^2) * cdf\left(\frac{\mu_4 - \mu_{max_{2,3}}}{\theta}\right)$$

$$+ \left(\mu_{max_{2,3}} + \mu_4\right) * \theta * pdf\left(\frac{\mu_{max_{2,3}} - \mu_4}{\theta}\right)$$

$$\mu_{max} = E[V_{max}]$$

$${\sigma^2}_{max} = E\left[{V_{max}}^2\right] - {\mu_{max}}^2$$

Clearly, this iterative process could be repeated indefinitely, allowing for a choice option set of arbitrarily large size.

Once the expected maximum value of the non-best options is determined, the DM will compare this value with the expected value of the best option. Thus, the decision variable (DV) will be a normally distributed random variable with mean equal to the difference of the best and max-non-best means and variance equal to the sum of the best and max-non-best variances:

$$DV \sim N\left(\Delta\mu, \sqrt{{\sigma_{best}}^2 + {\sigma_{max}}^2}\right)$$

$$\Delta\mu \triangleq \mu_{best} - \mu_{max}$$

Subjective uncertainty regarding option values eventually translates into uncertainty regarding what the "correct choice" is. This is summarized in terms of the probability P_c of "committing to the correct choice":

$$P_c = 1 - cdf_{DV}$$

$$\approx \frac{1}{1 + e^{\frac{\Delta\mu}{\sqrt{\frac{3}{\pi^2}({\sigma_{best}}^2 + {\sigma_{max}}^2)}}}}$$

where *cdf* is the cumulative density function of the corresponding normal distribution, evaluated at $DV = 0$, and the second line derives from a simple sigmoidal approximation to the normal cumulative density function (Daunizeau, 2017).

Acknowledgements

This research received funding from the European Union's Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreement Nos. 785907 and 945539 (Human Brain Project SGA2 and SGA3) to GP, the European Research Council under the Grant Agreement No. 820213 (ThinkAhead) to GP.

References

Balsdon, T., Wyart, V., & Mamassian, P. (2020). Confidence controls perceptual evidence accumulation. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-15561-w

Calder-Travis, J., Bogacz, R., & Yeung, N. (2020). *Bayesian confidence for drift diffusion observers in dynamic stimuli tasks*. https://doi.org/10.1101/2020.02.25.965384

Calder-Travis, J., Charles, L., Bogacz, R., & Yeung, N. (2016). *Bayesian confidence in optimal decisions*.

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. In *Unintended thought* (pp. 212–252). The Guilford Press.

Churchland, A. K., Kiani, R., & Shadlen, M. N. (2008). Decision-making with multiple alternatives. *Nat Neurosci*, *11*(6), 693–702. https://doi.org/10.1038/nn.2123

Cisek, P., Puskas, G. A., & El-Murr, S. (2009). Decisions in Changing Conditions: The Urgency-Gating Model. *The Journal of Neuroscience*.

Daunizeau, J. (2017). Semi-analytical approximations to statistical moments of sigmoid and softmax mappings of normal variables. *ArXiv:1703.00091 [q-Bio, Stat]*. http://arxiv.org/abs/1703.00091

De Martino, B., Fleming, S. M., Garrett, N., & Dolan, R. J. (2013). Confidence in value-based choice. *Nature Neuroscience*, *16*(1), 105–110. https://doi.org/10.1038/nn.3279

Ditterich, J. (2006). Evidence for time-variant decision making. *European Journal of Neuroscience*, *24*(12), 3628–3641. https://doi.org/10.1111/j.1460-9568.2006.05221.x

Drugowitsch, J., Mendonça, A. G., Mainen, Z. F., & Pouget, A. (2019). Learning optimal decisions with confidence. *Proceedings of the National Academy of Sciences*, *116*(49), 24872–24880. https://doi.org/10.1073/pnas.1906787116

Festinger, L. (1957). *A theory of cognitive dissonance* (pp. xi, 291). Stanford University Press.

Fleming, S. M., & Daw, N. D. (2017). Self-Evaluation of Decision-Making: A General Bayesian Framework for Metacognitive Computation. *Psychological Review*, *124*(1), 91–114. https://doi.org/10.1037/rev0000045

Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*(1594), 1338–1349. https://doi.org/10.1098/rstb.2011.0417

Fleming, S. M., Dolan, R. J., & Frith, C. D. (2012). Metacognition: Computation, biology and function. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1280–1286. https://doi.org/10.1098/rstb.2012.0021

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00443

Fontanesi, L., Gluth, S., Spektor, M. S., & Rieskamp, J. (2019). A reinforcement learning diffusion decision model for value-based decisions. *Psychonomic Bulletin and Review*, *26*, 1099–1121. https://doi.org/10.3758/s13423-018-1554-2

Forstmann, B. U., Ratcliff, R., & Wagenmakers, E.-J. (2016). Sequential Sampling Models in Cognitive Neuroscience: Advantages, Applications, and Extensions. *Annual Review of Psychology*, *67*(1), 641–666. https://doi.org/10.1146/annurev-psych-122414-033645

Fudenberg, D., Strack, P., & Strzalecki, T. (2018). Speed, Accuracy, and the Optimal Timing of Choices. *American Economic Review*, *108*(12), 3651–3684. https://doi.org/10.1257/aer.20150742

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annu Rev Neurosci*, *30*, 535–574. https://doi.org/10.1146/annurev.neuro.29.051605.113038

Gwinn, R., & Krajbich, I. (2020). Attitudes and attention. *Journal of Experimental Social Psychology*, *86*. https://doi.org/10.1016/j.jesp.2019.103892

Hauser, T. U., Moutoussis, M., Dayan, P., & Dolan, R. J. (2017). Increased decision thresholds trigger extended information gathering across the compulsivity spectrum. *Translational Psychiatry*, *7*(12), 1–10. https://doi.org/10.1038/s41398-017-0040-3

Hausmann-Thürig, D., & Läge, D. (2008). Sequential evidence accumulation in decision making: The individual desired level of confidence can explain the extent of information acquisition. *Judgment and Decision Making*, *3*(3), 229–243. https://doi.org/10.5167/uzh-6226

Kiani, R., & Shadlen, M. N. (2009). Representation of Confidence Associated with a Decision by Neurons in the Parietal Cortex. *Science*, *324*(5928), 759–764. https://doi.org/10.1126/science.1169405

Kvam, P. D. (2019). A geometric framework for modeling dynamic decisions among arbitrarily many alternatives. *Journal of Mathematical Psychology*, *91*, 14–37. https://doi.org/10.1016/j.jmp.2019.03.001

Kvam, P. D., & Pleskac, T. J. (2016). Strength and weight: The determinants of choice and confidence. *Cognition*, *152*, 170–180. https://doi.org/10.1016/j.cognition.2016.04.008

Lebreton, M., Abitbol, R., Daunizeau, J., & Pessiglione, M. (2015). Automatic integration of confidence in the brain valuation signal. *Nature Neuroscience*, *18*(8), 1159–1167. https://doi.org/10.1038/nn.4064

Lebreton, M., Jorge, S., Michel, V., Thirion, B., & Pessiglione, M. (2009). An Automatic Valuation System in the Human Brain: Evidence from Functional Neuroimaging. *Neuron*, *64*(3), 431–439. https://doi.org/10.1016/j.neuron.2009.09.040

Lee, D., & Daunizeau, J. (2020). Choosing what we like vs liking what we choose: How choice-induced preference change might actually be instrumental to decision-making. *PLoS ONE*, *15*(5). https://doi.org/10.1371/journal.pone.0231081

Lee, D., & Coricelli, G. (2020). An empirical test of the role of value certainty in decision making. *Frontiers in Psychology*, *11*. https://doi.org/10.1101/2020.06.16.155234

Lee, D., & Daunizeau, J. (2021). Trading Mental Effort for Confidence in the Metacognitive Control of Value-Based Decision-Making. b*ioRxiv*, 837054. https://doi.org/10.1101/837054

Lee, D., & Usher, M. (2020). Value Certainty in Drift-Diffusion Models of Preferential Choice. *bioRxiv*, 2020.08.22.262725. https://doi.org/10.1101/2020.08.22.262725

Li, H. H., & Ma, W. J. (2020). Confidence reports in decision-making with multiple alternatives violate the Bayesian confidence hypothesis. *Nature Communications*, *11*(1), 1–11. https://doi.org/10.1038/s41467-020-15581-6

Malhotra, G., Leslie, D. S., Ludwig, C. J. H., & Bogacz, R. (2018). Time-varying decision boundaries: Insights from optimality analysis. *Psychonomic Bulletin & Review*, *25*(3), 971–996. https://doi.org/10.3758/s13423-017-1340-6

Moran, R., Teodorescu, A. R., & Usher, M. (2015). Post choice information integration as a causal determinant of confidence: Novel data and a computational account. *Cognitive Psychology*, *78*, 99–147. https://doi.org/10.1016/j.cogpsych.2015.01.002

Moreno-Bote, R. (2010). Decision Confidence and Uncertainty in Diffusion Models with Partially Correlated Neuronal Integrators. *Neural Computation*, *22*(7), 1786–1811. https://doi.org/10.1162/neco.2010.12-08-930

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. https://doi.org/10.1109/TVLSI.2007.912191

Navarro-Martinez, D., Loomes, G., Isoni, A., Butler, D., & Alaoui, L. (2018). Boundedly rational expected utility theory. *Journal of Risk and Uncertainty*, *57*(3), 199–223. https://doi.org/10.1007/s11166-018-9293-3

Peixoto, D., Verhein, J. R., Kiani, R., Kao, J. C., Nuyujukian, P., Chandrasekaran, C., Brown, J., Fong, S., Ryu, S. I., Shenoy, K. V., & Newsome, W. T. (2021). Decoding and perturbing decision states in real time. *Nature*, 1–6. https://doi.org/10.1038/s41586-020-03181-9

Pleskac, T. J., & Busemeyer, J. R. (2010). *Two-Stage Dynamic Signal Detection: A Theory of Choice, Decision Time, and Confidence*. https://doi.org/10.1037/a0019737

Polanía, R., Woodford, M., & Ruff, C. C. (2019). Efficient coding of subjective value. *Nature Neuroscience*, *22*(1), 134–142. https://doi.org/10.1038/s41593-018-0292-0

Pouget, A., Drugowitsch, J., & Kepecs, A. (2016). Confidence and certainty: Distinct probabilistic quantities for different goals. *Nature Neuroscience*, *19*(3), 366–374. https://doi.org/10.1038/nn.4240

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion Decision Model: Current Issues and History. *Trends in Cognitive Sciences*, *20*(4), 260–281. https://doi.org/10.1016/j.tics.2016.01.007

Rollwage, M., Loosen, A., Hauser, T. U., Moran, R., Dolan, R. J., & Fleming, S. M. (2020). Confidence drives a neural confirmation bias. *Nature Communications*, *11*(1), 2634. https://doi.org/10.1038/s41467-020-16278-6

Roxin, A. (2019). Drift–diffusion models for multiple-alternative forced-choice decision making. *Journal of Mathematical Neuroscience*, *9*. https://doi.org/10.1186/s13408-019-0073-4

Schulz, L., Fleming, S. M., & Dayan, P. (2021). Metacognitive Computations for Information Search: Confidence in Control. *BioRxiv*, 2021.03.01.433342. https://doi.org/10.1101/2021.03.01.433342

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a Rational and Mechanistic Account of Mental Effort. *Annual Review of Neuroscience*, *40*(1), 99–124. https://doi.org/10.1146/annurev-neuro-072116-031526

Tajima, S., Drugowitsch, J., & Pouget, A. (2016). Optimal policy for value-based decision-making. *Nature Communications*, *7*, 12400. https://doi.org/10.1038/ncomms12400

Thura, D., & Cisek, P. (2014). Deliberation and Commitment in the Premotor and Primary Motor Cortex during Dynamic Decision Making. *Neuron*, *81*(6), 1401–1416. https://doi.org/10.1016/j.neuron.2014.01.031

Van Den Berg, R., Anandalingam, K., Zylberberg, A., Kiani, R., Shadlen, M. N., & Wolpert, D. M. (2016). A common mechanism underlies changes of mind about decisions and confidence. *ELife*, *5*(FEBRUARY2016), 1–21. https://doi.org/10.7554/eLife.12192

van den Berg, R., Zylberberg, A., Kiani, R., Shadlen, M. N., & Wolpert, D. M. (2016). Confidence Is the Bridge between Multi-stage Decisions. *Current Biology: CB*, *26*(23), 3157–3168. https://doi.org/10.1016/j.cub.2016.10.021

van Maanen, L., Fontanesi, L., Hawkins, G. E., & Forstmann, B. U. (2016). Striatal activation reflects urgency in perceptual decision making. *NeuroImage*, *139*, 294–303. https://doi.org/10.1016/j.neuroimage.2016.06.045

Voskuilen, C., Ratcliff, R., & Smith, P. L. (2016). Comparing fixed and collapsing boundary versions of the diffusion model. *Journal of Mathematical Psychology*, *73*, 59–79. https://doi.org/10.1016/j.jmp.2016.04.008

Wokke, M. E., Achoui, D., & Cleeremans, A. (2020). Action information contributes to metacognitive decision-making. *Scientific Reports*, *10*(1), 1–15. https://doi.org/10.1038/s41598-020-60382-y

Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: Confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1310–1321. https://doi.org/10.1098/rstb.2011.0416

Zenon, A., Solopchuk, O., & Pezzulo, G. (2019). An information-theoretic perspective on the costs of cognition. *Neuropsychologia*.

Zylberberg, A., Wolpert, D. M., & Shadlen, M. N. (2018). Counterfactual Reasoning Underlies the Learning of Priors in Decision Making. *Neuron*, *99*(5), 1083-1097.e6. https://doi.org/10.1016/j.neuron.2018.07.035