BRIEF REPORT

Early evidence affects later decisions: Why evidence accumulation is required to explain response time data

Jasper Winkel · Max C. Keuken · Leendert van Maanen ·
Eric-Jan Wagenmakers · Birte U. Forstmann

Abstract Models of decision making differ in how they treat early evidence as it recedes in time. Standard models, such as the drift diffusion model, assume that evidence is gradually accumulated until it reaches a boundary and a decision is initiated. One recent model, the urgency gating model, has proposed that decision making does not require the accumulation of evidence at all. Instead, accumulation could be replaced by a simple urgency factor that scales with time. To distinguish between these fundamentally different accounts of decision making, we performed an experiment in which we manipulated the presence, duration, and valence of early evidence. We simulated the associated response time and error rate predictions from the drift diffusion model and the urgency gating model, fitting the models to the empirical data. The drift diffusion model predicted that variations in the evidence presented early in the trial would affect decisions later in that same trial. The urgency gating model predicted that none of these variations would have any effect. The behavioral data showed clear effects of early evidence on the subsequent decisions, in a manner consistent with the drift diffusion model. Our results cannot be explained by the urgency gating model, and they provide support for an evidence accumulation account of perceptual decision making.

J. Winkel (✉) · M. C. Keuken · L. van Maanen ·
E.-J. Wagenmakers · B. U. Forstmann
Cognitive Science Center Amsterdam, University of Amsterdam,
Nieuwe Prinsengracht 130, 1018 VZ Amsterdam, The Netherlands
e-mail: jasperwinkel@gmail.com

M. C. Keuken · B. U. Forstmann
Max Planck Institute for Human Cognitive and Brain Sciences,
Leipzig, Germany

Keywords Decision making · Drift diffusion model · Urgency gating model · Evidence accumulation

All across the animal kingdom, survival directly depends on the adequacy of decision making under time pressure. For instance, animals often need to judge in a split second whether an approaching object is predator or prey. Mistaking predator for prey is costly, and hence it pays to analyze as much perceptual information as possible. On the other hand, the analysis of perceptual information takes valuable time—a correct classification (i.e., "yes, this is definitely a predator") is not worth much when one is in the process of being devoured. Hence, perceptual decision making involves a trade-off between speed and accuracy, and every decision is therefore determined by at least two factors: the quality of information that has been collected, and the level of caution displayed by the decision maker.

In order to understand and quantify the latent processes that drive simple decisions, several mathematical models have been formulated (Brown & Heathcote, 2008; Cisek, Puskas, & El-Murr, 2009; Ratcliff, 1978; Usher & McClelland, 2001). These models provide a formal account of how a decision is generated. In such models, evidence (such as perceptual information) is collected until a certain threshold is reached, indicating that enough confidence has been attained to commit to a particular alternative. Such accumulation models have proven to be instrumental in interpreting electrophysiological data related to the decision process (Churchland et al., 2011; Drugowitsch, Moreno-Bote, Churchland, Shadlen, & Pouget, 2012; Gold & Shadlen, 2007; Rangel & Hare, 2010). In addition to providing a framework for the interpretation of neuroscientific findings, the ability to estimate model parameters on an individual level also provides a powerful analysis tool for the study of both behavior and its neural correlates, at the group level, the individual level, or even the single-trial level (Forstmann et al., 2008; Mansfield,

Karayanidis, Jamadar, Heathcote, & Forstmann, 2011; Mulder, Wagenmakers, Ratcliff, Boekel, & Forstmann, 2012; Philiastides, Auksztulewicz, Heekeren, & Blankenburg, 2011; van Maanen et al., 2011; Wenzlaff, Bauer, Maess, & Heekeren, 2011; Winkel et al., 2012). Importantly, these models vary in a number of ways, including the influence of early evidence on later decisions. When fundamental differences exist between models, it is of great relevance to determine which model correctly represents the processes underlying decisions.

In most models of decision making, evidence is collected until it reaches a decision threshold, at which time a decision has been made. Sequential-sampling models, such as the drift diffusion model (DDM; Ratcliff, 1978) or the linear ballistic accumulator model (LBA; Brown & Heathcote, 2008), propose that all evidence is retained to inform the decision. The leaky competing accumulator model (Teodorescu & Usher, 2013; Usher & McClelland, 2001) includes a leak factor that causes early activation to decay over time. Another class of models implements a decision boundary that decreases with processing time (Ditterich, 2006; Drugowitsch et al., 2012). Despite their diversity, these models all share one assumption: Evidence is accumulated over time.

In stark contrast to such sequential-sampling models, however, one recent model has proposed that evidence need not be retained at all, but rather that the accumulation of evidence can be replaced entirely by an urgency factor. This urgency factor simply represents the increasing time pressure as the trial progresses, and it scales linearly with time (Cisek, Puskas, & El-Murr, 2009; see Eq. 2 in the Materials and Method section). Cisek et al. pointed out that under conditions in which evidence presentation is constant, a model with a multiplicative gain factor that increases linearly with time (i.e., an urgency factor) is mathematically equivalent to a model in which a constant level of evidence accumulates over time. From their behavioral data, Cisek et al. proposed that this urgency gating model (UGM) can account for the way that decisions are made better than various implementations of the DDM. The removal of the evidence accumulation process provides a fundamentally different account of decision making. The difference between the models can be clearly seen when evidence presentation is not constant during a trial. In that situation, the UGM and DDM predict different behavior. Most importantly, the DDM predicts that early evidence will also affect later decisions, whereas the UGM predicts that early evidence will be ignored if it is not immediately acted upon. This is illustrated in Fig. 1, which shows how constant or conflicting evidence is processed by the UGM, an accumulation model with incoming boundaries, and the DDM.

In contrast to the UGM's assumptions, previous experiments with monkeys have shown that early evidence can have profound effects on later decisions. In one experiment, a noisy stimulus was presented continuously (i.e., without obviously discrete trials) to each of two monkeys. A 100-ms burst of coherently moving dots was presented prior to the main

motion stimulus, influencing the monkeys' decisions 200 to 800 ms later (Huk & Shadlen, 2005). In another experiment, two monkeys had to withhold their responses during stimulus presentation, and coherent motion was present either at the beginning or the end of the presentation period. The results showed that motion energy early in the trial contributed more strongly toward the decision than did motion energy late in the trial (Kiani, Hanks, & Shadlen, 2008). The substantial impact of early motion energy is consistent with integration models of decision making and goes against leak-dominant models of decision making, in which the contribution of early evidence wanes as time progresses.

We designed an experiment to test the different predictions made by the DDM and the UGM, allowing us to adjudicate between these two qualitatively different models of decision making. We used a behavioral paradigm in which the amount of evidence changed over time, much like in the experiment by Cisek et al. (2009; see also Tsetsos, Chater, & Usher, 2012; Tsetsos, Gao, McClelland, & Usher, 2012; Tsetsos, Usher, & McClelland, 2011). One important difference, though, was that previously presented evidence did not remain visible throughout the trial. Using this experiment, we directly tested the predictions that early evidence would be incorporated in the later decision (DDM) or that it would be dismissed (UGM).

Materials and method

Participants

A group of 22 healthy participants (18 female, four male; mean age = 28.14 years, SD = 10.34) performed a calibration test and a random-dot-motion experiment during a single behavioral session. Three of the participants did not meet our inclusion criteria (i.e., >80% correct in the calibration task and <10% nonresponsive trials in the main task) and were excluded from further analyses. The experiment was approved by the local ethics committee of the University of Amsterdam. Participants were recruited from the University of Amsterdam student population and received either monetary compensation or research participation credits. The participants had normal or corrected-to-normal vision and did not have a history of neurological or psychiatric disorders, as assessed through self-report.

Stimuli

The experiment began with a calibration session to estimate each participant's sensitivity to the stimulus. This estimate determined the coherence (i.e., percentage of dots moving coherently) used in the experiment. Specifically, the coherence in the experiment was set so that it corresponded with an 80% accuracy level in the calibration session (as interpolated
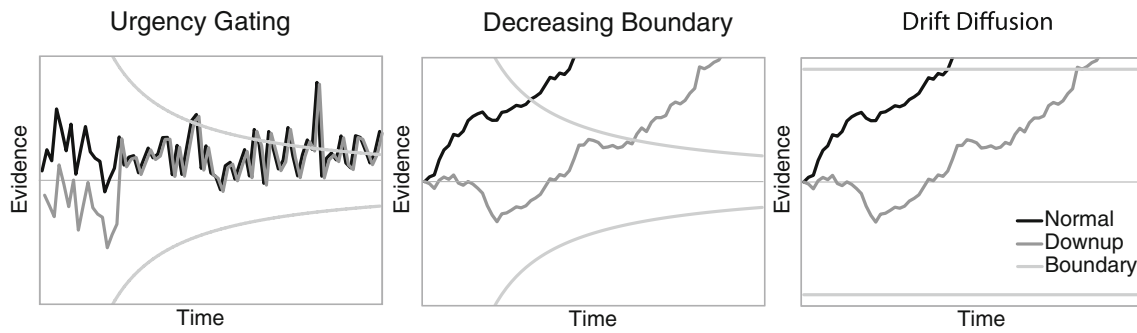
Fig. 1 Three ways to model decision processes. This figure illustrates the difference between modeling decision making as an urgency process, as accumulation with incoming boundaries, or as simple accumulation to a static boundary. The black and dark gray traces show the levels of evidence that each model represents during two sample trials. In the "normal" trace (black), evidence is directed upward continually. In the "down–up" trace (dark gray), evidence is directed first downward, and then upward. When the level of evidence meets the (light gray) decision threshold, a decision is made. The urgency gating model (UGM) has no accumulation factor and is driven by urgency only. The incoming-boundary model has both an accumulation and an urgency factor. The drift diffusion model has no urgency factor and is driven by accumulation only. For ease of comparison, we have visualized the gain parameter of the UGM as a decrease with time of the decision boundary. In the UGM, a decision is initiated when $tN \geq a$ (where N is the normal distribution representing the evidence, t is time, and a is threshold). With an incoming decision boundary, a decision is made when $N \geq a/t$, which is identical to the UGM decision rule (also see Drugowitsch et al., 2012). Note that the "down–up" trace in the left panel is shifted slightly in order to allow the overlapping traces both to be visible

from the psychometric curve produced by the proportional-rate diffusion model; Mulder et al., 2012; Palmer, Huk, & Shadlen, 2005). The calibration session was performed on a MacBook Pro (version 7.1, Apple), using custom software as well as the Psychophysics Toolbox for MATLAB (version 2007B, MathWorks). The actual behavioral task was a custom script implemented using the Presentation software (Neurobehavioral Systems), and it was presented on an LCD monitor with a 60-Hz refresh rate. In this task, participants applied continuous light pressure to two custom-built force sensors with their thumbs (van Campen, Keuken, van den Wildenberg, & Ridderinkhof, in press). Whenever participants released this baseline pressure, the experiment was paused. Responses were made by increasing the pressure on the thumb corresponding to the direction in which motion was perceived.

In the experiment (but not in the calibration session), we manipulated the course of evidence presentation within a trial, varying the presence or absence of early and late evidence, the correspondence of direction between the early and late evidence, and the length of the early evidence. When early evidence was presented, it lasted either 67 ms (short) or 117 ms (long), followed by a period with 0% coherence, the length of which was either 133 or 83 ms, chosen in such a way that the early evidence and the 0%-coherence sections always summed to 200 ms. When both early and late evidence was presented, the late evidence could be in the same direction as or the opposite direction from the early evidence. This resulted in nine different trial types. These trial types are described in full in Table 1, and the time course of stimulus presentation is illustrated in Fig. 2.

The experiment featured a total of 1,000 trials. Each trial was preceded by a 250-ms intertrial interval, during which a fixation cross was presented. During the trial, a random-dot-motion stimulus was presented for 2,000 ms or until the participant responded. Following the stimulus, a feedback message was presented for 500 ms. This feedback read "too fast" for RTs <200 ms and "too slow" for RTs >1,750 ms. If no response was given by 2,000 ms following stimulus onset, participants received the feedback "no response." Otherwise, the feedback message read "correct" or "incorrect." In trials on which late evidence was present, this feedback was determined on the basis of the late evidence. In trials without late evidence, the feedback was determined randomly, whereas the accuracy of the response was determined on the basis of the early evidence, for analysis purposes.

Participants were not informed that the evidence might change over the course of a trial. Following the experiment, we asked participants whether they had noticed a change in motion direction. The binary outcome of this exit interview was included as a between-subjects factor in the statistical analyses.

Table 1 Mean response times (RTs) and error rates across trial types

Trial Type	Early Evidence (0–200 ms)	Late Evidence (200–2000 ms)	Correct RT (SD)	Error Rate (SD)
1	0 ms	Absent	1,283 (97)	.50 (.08)
2	67 ms	Absent	1,268 (88)	.45 (.06)
3	117 ms	Absent	1,257 (75)	.43 (.07)
4	117 ms	Opposite	1,231 (68)	.19 (.08)
5	67 ms	Opposite	1,228 (60)	.17 (.08)
6	0 ms	Present	1,220 (66)	.16 (.08)
7	67 ms	Same	1,184 (75)	.12 (.07)
8	117 ms	Same	1,187 (70)	.12 (.08)
9	200 ms	Same	1,156 (67)	.09 (.06)

This table describes the different trial types and shows participants' mean RT for correct decisions, as well as their error rate, for each of these trial types. The trial types are ordered by the total level of evidence presented.
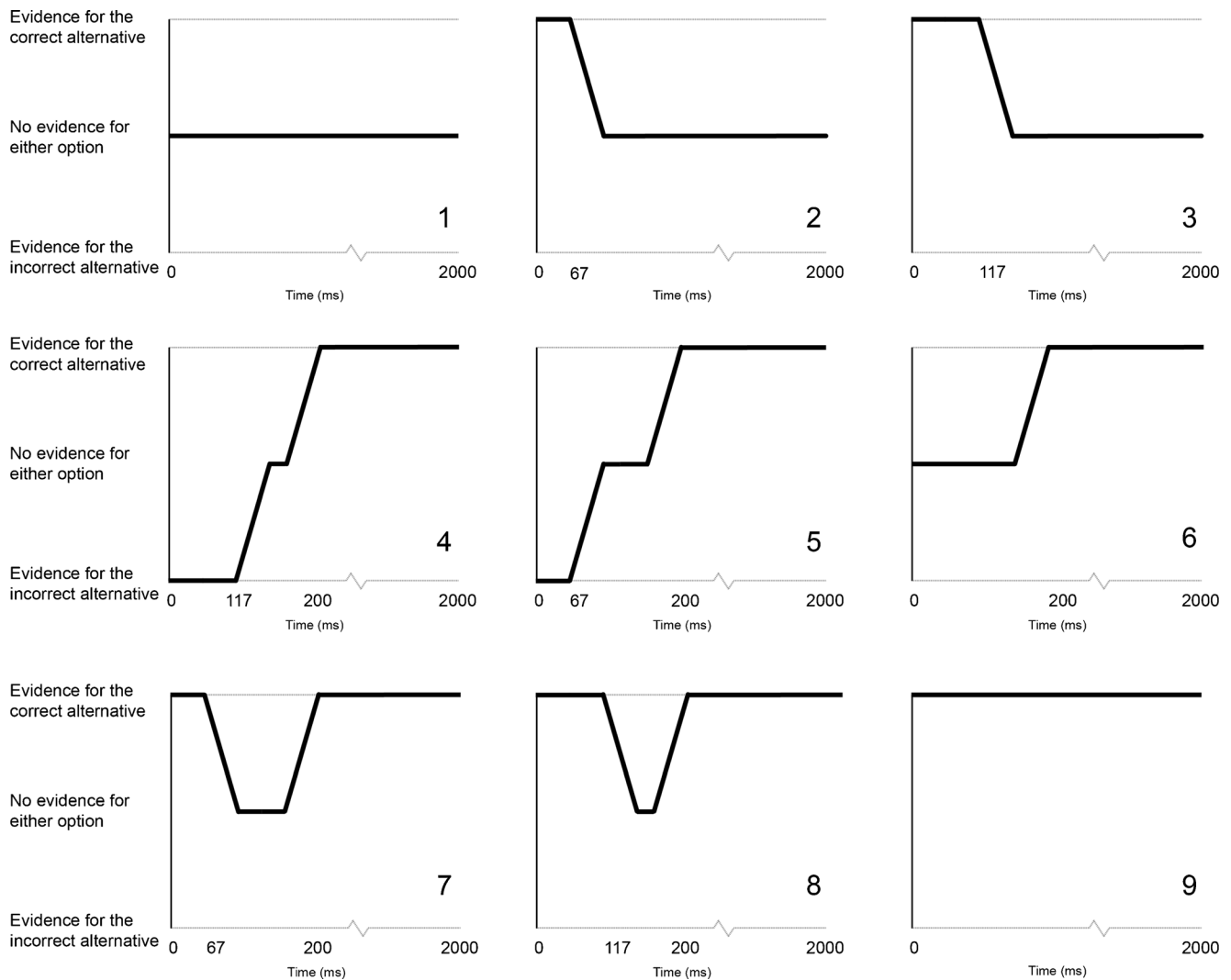
Fig. 2 Time courses of evidence presentation per trial types. This figure schematically illustrates how the presentation of evidence was varied between the nine different trial types. The top row shows the trial types without late information (1–3), whereas the next two rows show the trial types with late information (4–9). Note that the x-axis is discontinuous, greatly shortening the visual representation of the late evidence

Modeling

In order to show how the DDM and the UGM provide different predictions for within-trial variation of evidence, we fitted both models to the behavioral data (see Fig. 3). We produced trials from the "pure DDM"—that is, a DDM without trial-to-trial variability in model parameters (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006; Laming, 1968; see Eq. 1) and the UGM without temporal filtering (Cisek et al., 2009; see Eq. 2):

$$x(t + \Delta t) = x(t) + \mu(t)\Delta t + c\sqrt{\Delta t}, c \sim N(0, \sigma), x(0) = a/2 \quad (1)$$

and

$$x(t) = t[\mu(t) + c] + a/2, c \sim N(0, \sigma). \quad (2)$$

In these functions, x represents the current state of the decision variable, t is the time, the perceptual signal is drawn from input vector $\mu(t)$, and the noise c is drawn from a normal distribution with a mean of 0 and standard deviation σ. The input vector varies with time in a manner dependent on the trial type. In the DDM, the starting evidence $x(0)$ is set to $a/2$, so that there is no initial bias for one response over the other. A decision is initiated when $x \leq 0$ or $x \geq a$. The response time (RT) corresponding to that decision is the value of t at the time that the decision is initiated, plus a nondecision time t_0. We used a Δt of .001, corresponding to a time step size of 1 ms. We produced nine different input vectors, of which the timing and valence corresponded to those of the trial types used in the behavioral experiment. A SIMPLEX minimization routine (Nelder & Mead, 1965) was used to minimize the root-mean square error (RMSE) of the model predictions' mean RTs per
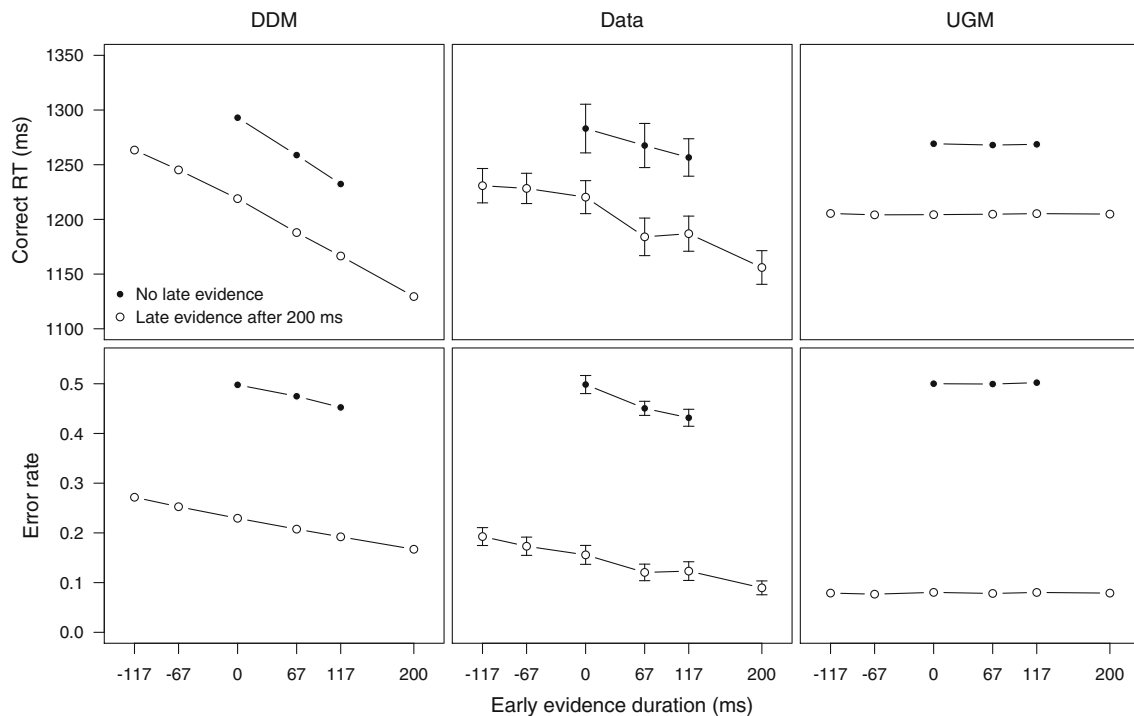
Fig. 3 Predicted and measured response times (RTs) and error rates. This figure shows the steady decrease in RTs and error rates as trial types contain more evidence for the correct response direction. The RT (top row) and error rate (bottom row) patterns are shown for the nine different trial types. The trial types without late evidence (filled circles) are plotted separately from those with late evidence (open circles). The x-axis represents the duration of the early evidence. If the early evidence points in the opposite direction from the late evidence (and thus, away from the correct response direction), it is given a negative value. The behavioral data are shown in the middle column, the drift diffusion model (DDM) fits are in the left column, and the urgency gating model (UGM) fits are in the right column. The error bars in the middle column correspond to the standard errors of the means

trial type and those of the behavioral data. Although the models were fit to the mean RT values of the different trial types, we let the models produce the corresponding accuracy levels freely.

We constrained the models in two ways. Models that produced >15% of RTs >2,000 ms were penalized by assigning an RMSE value of 10^{10} ms, as were models that produced >5% of decision times <117 ms.

The complete R code for our simulations and fits is available as supplementary material to this article.

Analyses

We performed four repeated measures analyses of variance. We analyzed RTs and error rates separately as dependent measures, while using Trial Type as a within-subjects factor. This was repeated separately for the three trial types without late evidence and for the six trial types with late evidence. Consistent with our modeling outcomes (see the Results—Modeling section, below), we included a polynomial contrast to estimate the linear component of trial type, ordered by the total amount of evidence presented. In addition, we performed the same analyses separately on the three trial types with no late evidence (in

which only noise was presented in the later part of the stimulus) and the six trial types with late evidence (in which the later part of the stimulus contained coherent motion). For all of these analyses, we report degrees of freedom and p values following the Greenhouse–Geisser correction for nonsphericity.

Results

Modeling

The DDM provided a better fit to the data (RMSE = 15 ms) than the UGM did (RMSE = 23 ms). More importantly, the DDM matched the qualitative pattern across trial types, which the UGM does not (see Fig. 3). Although both models predicted differences between the trial types with and without late evidence, the DDM predicted effects of the presence, duration, and direction of the early evidence that the UGM did not.

Behavioral data

As can be seen in the upper middle panel of Fig. 3, RTs continuously decreased as the trial types contained more

consistent evidence. The analysis of the three trial types without late evidence showed no significant difference [main effect, $F(1, 1.8) = 2.0$, $p = .15$; linear contrast, $F(1, 1) = 3.7$, $p = .07$]. The six trial types with late evidence still showed a continuous decrease in RTs with increasing early evidence [$F(1, 2.7) = 35.7$, $p < .001$; linear contrast, $F(1, 1) = 164.6$, $p < .001$], corresponding to the pattern predicted by the DDM.

Similar to RTs, the bottom row of Fig. 3 shows that error rates continuously decreased with increasing early evidence. This effect was found in the analysis of the three trial types without late evidence [main effect, $F(1, 1.7) = 4.4$, $p = .027$; linear contrast, $F(1, 1) = 6.2$, $p = .022$], as well as in the six trial types with late evidence [main effect, $F(1, 3.4) = 20.0$, $p < .001$; linear contrast, $F(1, 1) = 50.6$, $p < .001$], again corresponding to the pattern predicted by the DDM.

Exit interview

Out of the 19 participants, seven reported that they saw a change in motion direction. Including this measure of awareness as a between-subjects factor did not change the pattern of results. In addition, the awareness factor was not statistically significant in any analysis of RTs or error rates. The results reported above represent analyses including this factor.

Discussion

In this experiment, we tested the predictions from the drift diffusion model and the urgency gating model. The DDM assumes that evidence is accumulated over time. This leads to the prediction that increasing levels of early evidence will cause later decisions to be both faster and more accurate. The UGM assumes that evidence is not accumulated. The leads to the prediction that trial types that have the same late evidence should have the same speed and accuracy, regardless of early information. Our behavioral data clearly showed that both speed and accuracy improved with the presence of early evidence. Consequently, the specific predictions made by the UGM were falsified by our behavioral data. Our data are qualitatively consistent with those predicted by the DDM. Together with earlier findings that also showed a strong influence of early evidence on later decisions (Huk & Shadlen, 2005; Kiani et al., 2008), these findings strongly suggest that an accurate account of perceptual decision making requires the gradual accumulation of evidence.

What remains is to explain the findings in Cisek et al. (2009), suggesting that the UGM can account for the behavioral data better than the DDM. Critically, the paradigm used in the Cisek et al. study differed from the paradigms commonly used in perceptual decision making (Mulder et al., 2012; Roitman & Shadlen, 2002; Ruff, Marrett, Heekeren, Bandettini, & Ungerleider, 2010; van Maanen et al., 2011;

Winkel et al., 2012; see Bogacz, Wagenmakers, Forstmann, & Nieuwenhuis, 2010; Gold & Shadlen, 2007; Heekeren, Marrett, & Ungerleider, 2008, for reviews) in an important way. Specifically, in the Cisek et al. paradigm, all evidence presented remained visible throughout the entire trial, resulting in a stimulus that displayed the integral of the presented evidence. This integral was then used as the "current level of evidence," which was provided as the input evidence to both the UGM and the DDM. By providing the models with the integral of the evidence, Cisek et al. effectively used the UGM as an accumulator model with an added urgency term. Although this is not problematic in itself, it does not support the claim that urgency can replace accumulation, as was originally proposed by Cisek et al.

Since its original publication, the UGM has been adapted to account for more continuous evidence presentation (Thura, Beauregard-Racine, Fradet, & Cisek, 2012). This adaptation no longer uses the level of current information directly, but rather integrates the history of changes in evidence strength. Thus, the adapted version of the model can predict behavior in a random-dot-motion task, because the accumulated changes in evidence strength can now be represented as the net motion direction, rather than as a visibly represented number of dots. However, the revised model still predicts that neither the duration nor the valence of early evidence should contribute to a later decision, if they do not directly lead to a decision. The behavioral findings published along with these model revisions do indeed lack any effect of the variation in early evidence, which is in direct contradiction with our present findings.

Although the UGM proposes that all previous information is forgotten, other models, such as the LCA, propose that the represented evidence level gradually decays. Our data do not allow us to rule out the presence of a small leak. However, a large leak seems unlikely, considering that the first 67 ms of evidence have pronounced effects 1,100 ms later. Both the optimal level of leak and the level of leak exhibited by human participants vary depending on task contingencies (see, e.g., Ossmy et al., 2013).

Although our data showed that urgency cannot replace accumulation, we do not wish to deny that urgency can sometimes play an important role in decision making. It is quite possible that a constant decision threshold such as the one assumed by the DDM cannot accurately account for all of the observed data in decision-making experiments. Evidence for this notion comes from a time-variant effect of prior evidence (Hanks, Mazurek, Kiani, Hopp, & Shadlen, 2011; but see van Ravenzwaaij, Mulder, Tuerlinckx, & Wagenmakers, 2012). Ditterich (2006) also argued for a gain that increased over time during the decision process. It is important to note that unlike Cisek et al. (2009), Ditterich does not propose to replace evidence accumulation with this gain. Ditterich's model scales presented evidence with a gain

factor that increases over time, but it still accumulates resulting evidence in a manner similar to the DDM. The addition of an urgency factor makes sense from an optimal decision-making standpoint, if a cost is associated with additional observation (Drugowitsch et al., 2012). This optimality assumes that the quality of the incoming evidence is known. If it is unknown, the shape of the decision boundary should be determined by the quality of the evidence sampled so far (Deneve, 2012). Note that, although all three of these models argue against a static decision boundary, they differ fundamentally from the UGM in that they supplement evidence accumulation with urgency, rather than replacing it. Incoming boundaries may be needed to explain why, in Cisek et al.'s study, decisions were made at different total evidence levels (see their Fig. 7, column 3). Such incoming bounds are also consistent with findings of increased overall activation levels across groups of neurons with different response selectivities (Churchland, Kiani, & Shadlen, 2008; Drugowitsch et al., 2012).

One reason why it is important to correctly formulate decision-making models is that they help to assign a functional role to activity found in electrophysiological experiments (Gold & Shadlen, 2007; Rangel & Hare, 2010). One example of this concerns the seminal electrophysiological findings by Shadlen and Newsome (2001), demonstrating that during decision making, directionally sensitive neurons in the lateral intraparietal area of the macaque show a steady increase in firing toward a fixed threshold (Roitman & Shadlen, 2002; Shadlen & Newsome, 2001). Cisek et al. (2009) interpreted these finding not as representing an increasing level of gathered evidence, but rather as a static level of evidence combined with an increasing level of urgency. This interpretation has since been acknowledged (Domenech & Dreher, 2010; Simen, 2012; van Vugt, Simen, Nystrom, Holmes, & Cohen, 2012; Zhang, 2012) and supported (Rangel & Hare, 2010; Standage, You, Wang, & Dorris, 2011). The feasibility of such an interpretation has recently been tested by examining its predictions for the correlation and variance structure of the neural firing rates in lateral intraparietal area across time (Churchland et al., 2011). As in the present article, the results of that study are consistent with evidence accumulation, rather than a time-dependent scaling of the currently presented evidence.

An important future question is whether the addition of an urgency factor accounts for sufficient additional variance to warrant its inclusion. The answer to this question can be found using model comparison methods, and it may not universally apply to all tasks and participants. Although an urgency factor is a plausible mechanism for decision making, its necessity may well depend on task parameters such as response deadlines, reward rates, and intertrial intervals, as well as individual or species differences. Much recent discussion has focused on the shape of the decision boundary. Studies have provided

conflicting answers, suggesting either that an additional urgency parameter is warranted (Drugowitsch et al., 2012) or that it is not (Milosavljevic, Malmaud, Huth, Koch, & Rangel, 2010). Although our experiment does not directly address that issue, it does demonstrate that urgency cannot replace accumulation when modeling decision making.

References

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,* 700–765. doi:10.1037/0033-295X.113.4.700

Bogacz, R., Wagenmakers, E.-J., Forstmann, B. U., & Nieuwenhuis, S. (2010). The neural basis of the speed–accuracy tradeoff. *Trends in Neurosciences, 33,* 10–16. doi:10.1016/j.tins.2009.09.002

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology, 57,* 153–178. doi:10.1016/j.cogpsych.2007.12.002

Churchland, A. K., Kiani, R., Chaudhuri, R., Wang, X.-J., Pouget, A., & Shadlen, M. N. (2011). Variance as a signature of neural computations during decision making. *Neuron, 69,* 818–831. doi:10.1016/j.neuron.2010.12.037

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

Cisek, P., Puskas, G. A., & El-Murr, S. (2009). Decisions in changing conditions: The urgency-gating model. *Journal of Neuroscience, 29,* 11560–11571. doi:10.1523/JNEUROSCI.1844-09.2009

Deneve, S. (2012). Making decisions with unknown sensory reliability. *Frontiers in Decision Neuroscience, 6,* 75. doi:10.3389/fnins.2012.00075

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

Domenech, P., & Dreher, J.-C. (2010). Decision threshold modulation in the human brain. *Journal of Neuroscience, 30,* 14305–14317. doi:10.1523/JNEUROSCI.2371-10.2010

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,* 3612–3628. doi:10.1523/JNEUROSCI.4010-11.2012

Forstmann, B. U., Dutilh, G., Brown, S., Neumann, J., von Cramon, D. Y., Ridderinkhof, K. R., & Wagenmakers, E.-J. (2008). Striatum and pre-SMA facilitate decision-making under time pressure. *Proceedings of the National Academy of Sciences, 105,* 17538–17542. doi:10.1073/pnas.0805903105

Forstmann, B. U., Brown, S., Dutilh, G., Neumann, J., & Wagenmakers, E.-J. (2010). The neural substrate of prior information in perceptual decision making: A model-based analysis. *Frontiers in Human Neuroscience, 4,* 40. doi:10.3389/fnhum.2010.00040

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience, 30,* 535–574. doi:10.1146/annurev.neuro.29.051605.113038

Hanks, T. D., Mazurek, M. E., Kiani, R., Hopp, E., & Shadlen, M. N. (2011). Elapsed decision time affects the weighting of prior probability in a perceptual decision task. *Journal of Neuroscience, 31,* 6339–6352. doi:10.1523/JNEUROSCI.5613-10.2011

Heekeren, H. R., Marrett, S., & Ungerleider, L. G. (2008). The neural systems that mediate human perceptual decision making. *Nature Reviews Neuroscience, 9,* 467–479. doi:10.1038/nrn2374

Huk, A. C., & Shadlen, M. N. (2005). Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *Journal of Neuroscience, 25,* 10420–10436. doi:10.1523/JNEUROSCI.4684-04.2005

Kiani, R., Hanks, T. D., & Shadlen, M. N. (2008). Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *Journal of Neuroscience, 28,* 3017–3029. doi:10.1523/JNEUROSCI.4761-07.2008

Laming, D. R. J. (1968). *Information theory of choice-reaction times*. London, UK: Academic Press.

Mansfield, E. L., Karayanidis, F., Jamadar, S., Heathcote, A., & Forstmann, B. U. (2011). Adjustments of response threshold during task switching: A model-based functional magnetic resonance imaging study. *Journal of Neuroscience, 31,* 14688–14692. doi:10.1523/JNEUROSCI.2390-11.2011

Milosavljevic, M., Malmaud, J., Huth, A., Koch, C., & Rangel, A. (2010). The Drift Diffusion Model can account for the accuracy and reaction time of value-based choices under high and low time pressure. *Judgment and Decision Making, 5,* 437–449. doi:10.2139/ssrn.1901533

Mulder, M. J., Wagenmakers, E.-J., Ratcliff, R., Boekel, W., & Forstmann, B. U. (2012). Bias in the brain: A diffusion model analysis of prior probability and potential payoff. *Journal of Neuroscience, 32,* 2335–2343. doi:10.1523/JNEUROSCI.4156-11.2012

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal, 7,* 308–313. doi:10.1093/comjnl/7.4.308

Ossmy, O., Moran, R., Pfeffer, T., Tsetsos, K., Usher, M., & Donner, T. H. (2013). The timescale of perceptual evidence integration can be adapted to the environment. *Current Biology, 23,* 1–6. doi:10.1016/j.cub.2013.04.039

Palmer, J., Huk, A. C., & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of Vision, 5*(5):1, 376–404. doi:10.1167/5.5.1

Philiastides, M. G., Auksztulewicz, R., Heekeren, H. R., & Blankenburg, F. (2011). Causal role of dorsolateral prefrontal cortex in human perceptual decision making. *Current Biology, 21,* 980–983. doi:10.1016/j.cub.2011.04.034

Rangel, A., & Hare, T. (2010). Neural computations associated with goal-directed choice. *Current Opinion in Neurobiology, 20,* 262–70. doi:10.1016/j.conb.2010.03.001

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108. doi:10.1037/0033-295X.85.2.59

Roitman, J. D., & Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of Neuroscience, 22,* 9475–9489.

Ruff, D. A., Marrett, S., Heekeren, H. R., Bandettini, P. A., & Ungerleider, L. G. (2010). Complementary roles of systems representing sensory evidence and systems detecting task difficulty during perceptual decision making. *Frontiers in Decision Neuroscience, 4,* 190. doi:10.3389/fnins.2010.00190

Shadlen, M. N., & Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of Neurophysiology, 86,* 1916–1936.

Simen, P. (2012). Evidence accumulator or decision threshold—Which cortical mechanism are we observing? *Frontiers in Psychology, 3,* 183. doi:10.3389/fpsyg.2012.00183

Standage, D., You, H., Wang, D.-H., & Dorris, M. C. (2011). Gain modulation by an urgency signal controls the speed–accuracy trade-off in a network model of a cortical decision circuit. *Frontiers in Computational Neuroscience, 5,* 7. doi:10.3389/fncom.2011.00007

Teodorescu, A. R., & Usher, M. (2013). Disentangling decision models: From independence to competition. *Psychological Review, 120,* 1–38. doi:10.1037/a0030776

Thura, D., Beauregard-Racine, J., Fradet, C.-W., & Cisek, P. (2012). Decision making by urgency gating: Theory and experimental support. *Journal of Neurophysiology, 108,* 2912–2930. doi:10.1152/jn.01071.2011

Tsetsos, K., Chater, N., & Usher, M. (2012a). Salience driven value integration explains decision biases and preference reversal. *Proceedings of the National Academy of Sciences, 109,* 9659–9664. doi:10.1073/pnas.1119569109

Tsetsos, K., Gao, J., McClelland, J. L., & Usher, M. (2012b). *Using time-varying evidence to test models of decision dynamics: bounded diffusion vs. the leaky competing accumulator model. Frontiers in Decision Neuroscience, 6,* 79. doi:10.3389/fnins.2012.00079

Tsetsos, K., Usher, M., & McClelland, J. L. (2011). Testing multi-alternative decision models with non-stationary evidence. *Frontiers in Decision Neuroscience, 5,* 63. doi:10.3389/fnins.2011.00063

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review, 108,* 550–592. doi:10.1037/0033-295X.111.3.757

van Campen, A. D., Keuken, M. C., van den Wildenberg, W. P. M., & Ridderinkhof, K. R. (in press). TMS over M1 reveals expression and selective suppression of conflicting action impulses. *Journal of Cognitive Neuroscience*. doi:10.1162/jocn_a_00482

van Maanen, L., Brown, S. D., Eichele, T., Wagenmakers, E.-J., Ho, T., Serences, J., & Forstmann, B. U. (2011). Neural correlates of trial-to-trial fluctuations in response caution. *Journal of Neuroscience, 31,* 17488–17495. doi:10.1523/JNEUROSCI.2924-11.2011

van Ravenzwaaij, D., Mulder, M. J., Tuerlinckx, F., & Wagenmakers, E.-J. (2012). Do the dynamics of prior information depend on task context? An analysis of optimal performance and an empirical test. *Frontiers in Psychology, 3,* 132. doi:10.3389/fpsyg.2012.00132

van Vugt, M. K., Simen, P., Nystrom, L. E., Holmes, P., & Cohen, J. D. (2012). EEG oscillations reveal neural correlates of evidence accumulation. *Frontiers in Decision Neuroscience, 6,* 106. doi:10.3389/fnins.2012.00106

Wenzlaff, H., Bauer, M., Maess, B., & Heekeren, H. R. (2011). Neural characterization of the speed–accuracy tradeoff in a perceptual decision-making task. *Journal of Neuroscience, 31,* 1254–1266. doi:10.1523/JNEUROSCI.4000-10.2011

Winkel, J., van Maanen, L., Ratcliff, R., van der Schaaf, M. E., van Schouwenburg, M. R., Cools, R., & Forstmann, B. U. (2012). Bromocriptine does not alter speed–accuracy tradeoff. *Frontiers in Decision Neuroscience, 6,* 126. doi:10.3389/fnins.2012.00126

Zhang, J. (2012). The effects of evidence bounds on decision-making: Theoretical and empirical developments. *Frontiers in Psychology, 3,* 263. doi:10.3389/fpsyg.2012.00263