Cognition 193 (2019) 104022

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Brief article

Integration to boundary in decisions between numerical sequences

Moshe Glickman[a,*], Marius Usher[a,b,*]

[a] School of Psychology, University of Tel Aviv, Israel
[b] Sagol School of Neuroscience, University of Tel Aviv, Israel

ARTICLE INFO

Keywords:
Integration-to-boundary
Fixed boundaries
Collapsing boundaries
Adaptation
Numerical cognition
Decision strategies

ABSTRACT

Integration-to-boundary is a prominent normative principle used in evidence-based decisions to explain the speed-accuracy trade-off and determine the decision-time. Despite its prominence, however, the decision boundary is not directly observed, but rather is theoretically assumed, and there is still an ongoing debate regarding its form: fixed vs. collapsing. The aim of this study is to show that the integration-to-boundary process extends to decisions between rapid pairs of numerical sequences (2 Hz rate), and to determine the boundary type by directly monitoring the noisy accumulated evidence. In a set of two experiments (supplemented by computational modelling), we demonstrate that integration to a collapsing-boundary takes place in such tasks, ruling out non-integration heuristic strategies. Moreover, we show that participants can adaptively adjust their boundaries in response to reward contingencies. Finally, we discuss the implications to decision optimality and the nature of processes and representations in numerical cognition.

1. Introduction

Effective decision-making requires careful balancing between the cost of deliberation time and the quality of the decision. A parsimonious algorithm thought to minimize the time needed to achieve a specified accuracy is *integration-to-boundary* (Ratcliff, Smith, Brown, & McKoon, 2016; Wald, 1947). This algorithm, first deployed by Alan Turing in the decoding of the *Enigma* code (Gold & Shadlen, 2002), is now the default mechanism assumed to operate in a wide range of decisions, ranging from perceptual or lexical decisions (Ratcliff & Rouder, 1998; Smith, Ratcliff, & Sewell, 2014) to memory (Ratcliff, 1978) and even choices between food items (Krajbich, Armel, & Rangel, 2010). Its essence is the integration of noisy samples towards a response-boundary, whose variation accounts for the speed-accuracy trade-off. Yet, despite its prominence, in most behavioral studies the integration of noisy-samples, as well as the *decision-boundary*, remain theoretical entities that are not directly measured or controlled (but see Malhotra, Leslie, Ludwig, & Bogacz, 2017). While the decision boundary can be monitored in physiological studies, there is debate about its functional time-invariance (Drugowitsch, Moreno-Bote, Churchland, Shadlen, & Pouget, 2012; Hawkins, Forstmann, Wagenmakers, Ratcliff, & Brown, 2015).

One important type of decision, ubiquitous in day-to-day life, involves sequences of numerical values (or payoffs), such as choosing stocks on the basis of past returns or selecting a hotel on the basis of online rating. Prominent theories in numerical cognition have proposed that symbolic numerical values are associated with representations of magnitudes (Dehaene, 2011; Dehaene, Molko, Cohen, & Wilson, 2004). Thus, an interesting possibility is that numerical representations, like perceptual samples of evidence, are subject to evidence-integration. While a number of studies have examined choices between rapid sequences of numerical payoffs (Brusovansky, Vanunu, & Usher, 2017; Glickman, Tsetsos, & Usher, 2018; Spitzer, Waschke, & Summerfield, 2017; Tsetsos, Chater, & Usher, 2012; Vanunu, Pachur, & Usher, 2018) or numerosity displays (Zeigenfuse, Pleskac, & Liu, 2014), these studies could not establish integration-to-boundary, as they did not use a free-response paradigm that enables the subject control over the stopping-time.

The aim of this study is threefold. First, we aimed to demonstrate the normative integratio!n-to-boundary in a design that allows us control over the noisy-samples, and thus to obtain a more direct measure of the decision-boundary. Specifically, we examined the support for a fixed or a collapsing boundary (Hawkins et al., 2015). Second, we probed if integration-to-boundary takes place outside the perceptual domain, particularly in decisions between pairs of rapid numerical sequences. This is not an obvious generalization, as numbers are typically subject to symbolic computations, which are consistent with simple heuristics (e.g., a cut-off to decide at the first value that exceeds it, or counting the number of “pair-winners“ in favor of each alternative;

* Corresponding authors at: The School of Psychological Sciences, Tel Aviv University, Ramat Aviv, POB 39040, Tel Aviv 69978, Israel (M. Glickman). The School of Psychological Sciences and the Sagol School of Neuroscience, Tel Aviv University, Ramat Aviv, POB 39040, Tel Aviv 69978, Israel (M. Usher).

E-mail addresses: glickman@mail.tau.ac.il (M. Glickman), marius@post.tau.ac.il (M. Usher).

https://doi.org/10.1016/j.cognition.2019.104022
Received 31 December 2018; Received in revised form 3 June 2019; Accepted 4 July 2019

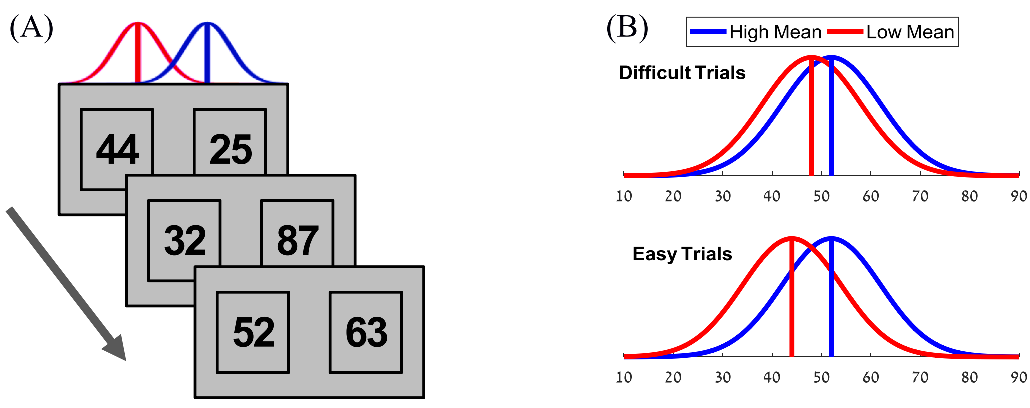

Fig. 1. Experimental paradigm. (A) The sequences of two-digit numbers, selected from Gaussian distributions are presented at a rate of 2 pairs/sec, and till the subject responds (left/right) to indicate the alternative that is larger on average. The sequences and the decisions participants make in each trial are recorded. (B) The difficulty manipulation involved a shift in the mean of the lower distribution (red curves); the larger-mean distribution (blue curves) remained the same. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 1A). Furthermore, we contrasted between different types of evidence-integration mechanisms (independent vs. competitive; Teodorescu & Usher, 2013) or integration to a decision-boundary vs. integration without boundary (using a random-timer for decision termination). Third, we asked if subjects can adaptively control the decision-boundary in response to task demands. To answer these questions, we conducted two experiments, in which participants were presented with choices between rapid numerical sequences (Fig. 1A), which are terminated by the subject's decision; in Exp. 2 we also manipulated (within participants) the time-cost of the decision.

To preview our results, we find that integration to a collapsing-boundary (Malhotra et al., 2017; Palestro, Weichart, Sederberg, & Turner, 2018) accounts best to choice and decision-time data, and that the pay-off manipulation indeed affects the shape of the collapsing-boundary. We start by presenting a computational study that lays out the predictions of the integration-to-boundary algorithm (either fixed or collapsing), and distinguishes it from alternative non-integration to boundary models. Following, we present our experiments, and carry out computational modeling that validates our conclusion by accounting for response-times (RT) distributions.

2. Computational predictions

We contrasted four choice models: (i) integration to boundary (fixed/collapsing) with either high or low noise, (ii) timer model, (iii) heuristics based on the first value or difference exceeding a predefined cut-off, and (iv) counting heuristics. Note that, unlike heuristics which are based on symbolic computations, the integration-to-boundary model assumes two distinct sources of noise: *external-noise*, corresponding to variability in the samples (Fig. 1), and *internal-noise*, corresponding to the variability in the encoding of the numbers into a magnitude representation during the evidence-integration.

We simulated the models using the following assumptions:

(i) *Integration-to-boundary*. We employed a "pure" diffusion model (without additional variability parameters; Mazurek, Roitman, Ditterich, & Shadlen, 2003; Wagenmakers, Van Der Maas, & Grasman, 2007; but see the Results of Exp. 1 for a comparison involving a diffusion model which includes between trial variability parameters; Ratcliff & McKoon, 2008). We used the following difference equations:

$$X(t) = X(t-1) + \mu(t) + \varepsilon(t),\ \varepsilon \sim N(0, \sigma_{internal}{}^2) \quad (1)$$

where $X(t)$ is the accumulated differences between the sequences at time t, $\mu(t)$ is the difference between the two samples at time t (note that this includes the external/sampling noise), and $\varepsilon(t)$ is a random internal noise sampled from $N(0, \sigma_{internal}{}^2)$. Decisions were made when $X(t)$ exceeds one of two symmetrical boundaries, $\pm u(t)$. We tested two types of boundaries: (i) fixed boundaries, $u(t) = c$, where c is a constant, (ii) collapsing boundaries, modeled using a Weibull cumulative distribution function (Hawkins et al., 2015):

$$u(t) = a - \left[1 - \exp\left(-\left(\frac{t}{\lambda}\right)^k\right)\right]\cdot(a - a') \quad (2)$$

where $\pm u(t)$ are the upper/lower thresholds at time t, a/a' are the initial (intercept) and asymptotic values of the boundary, respectively, and λ/k are the scale and shape parameters of the Weibull function, respectively. We simulated both types of models 100,000 times, using samples drawn from Gaussians with $\mu_1 = 52$, $\mu_2 = 46$, and $\sigma_{external} = 10$, and the following set of parameters for the fixed-boundary model: $c = 35$, $\sigma_{internal} = 10$ (high noise) or $\sigma_{internal} = 0.5$ (low noise), and the following ones for the collapsing-boundary model: $a = 60$, $k = 3$, $\lambda = 5.5$, $a' = 20$.[1] The integration process was terminated once the integrated evidence (including internal noise) exceeds one of the boundaries.

(ii) *Timer model.* The integration process was similar to the one described in the *Integration-to-boundary* section. However, the stopping-time was randomly sampled from an ex-Gaussian distribution ($\mu = 3$, $\sigma = 0.5$ and $\lambda = \frac{2}{3}$), reflecting a process that is exogenous to the integration of evidence.

(iii) *Cut-off heuristics.* We examined two types of cut-off heuristics, both do not assume integration of evidence: (i) value cut-off - observers choose the sequence in which the first number exceeds a predetermined threshold, (ii) differences cut-off - observers choose based on the first frame in which the difference between the numbers exceeds a predetermined threshold. We simulated each model 100,000 times using cut-offs of 60 and 70 for the value cut-off heuristic, and 20 and 30 for the differences cut-off heuristic.

(iv) *Counting heuristic.* Decisions are based on tallying the number of pair-winners of each alternative (input of +1 for the pair-winner and 0 for the pair-loser). Decision is made when one of the tallies exceeds a predetermined threshold (absolute criterion), or alternatively, when the difference in tallies exceeds a predetermined threshold (relative criterion).

For each integration model, we examined how the integrated-evidence (including the external/sampling noise but excluding the internal-noise) at the time of response (i.e., the cumulative sum of differences between the sequences) varies as a function of decision-time for the correct responses. Note that this (external) integrated-evidence is a variable which we can measure in each trial (as we have access to the samples), and which does not have to be identical to the boundary (because we excluded the internal noise, to which we do not have

[1] All parameters used in this section, were selected to keep the accuracy and decision-time in the experimental range.

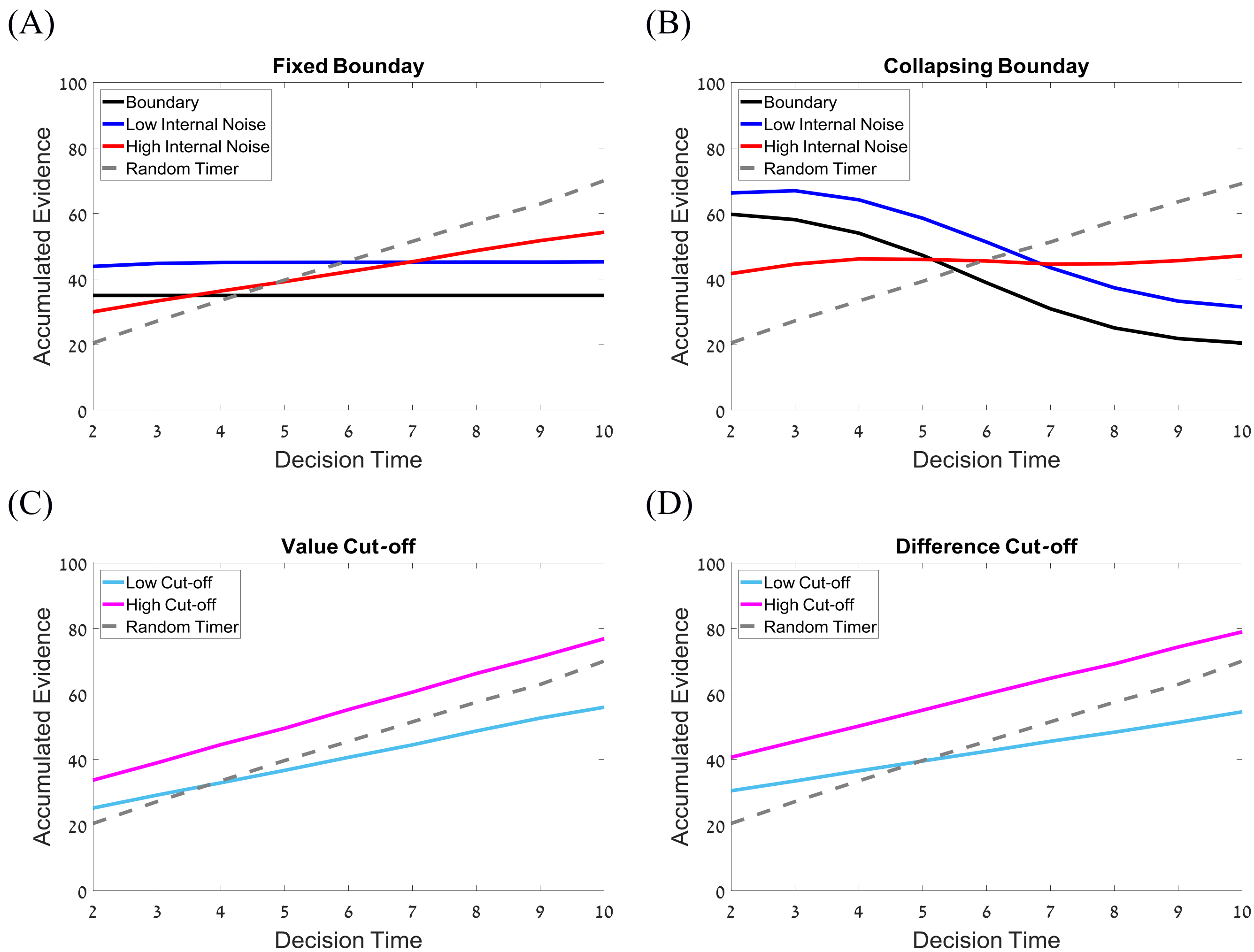

Fig. 2. Models predictions. Predictions for the fixed integration-to-boundary model are shown in (A), and for the collapsing integration-to-boundary model are shown in (B). The boundaries are shown with solid black-lines. In each panel, the (external) integrated-evidence model predictions are shown for low internal noise (blue lines) and for high internal-noise (red lines). The random-timer model (dashed gray line) predicts increase of the mean accumulated evidence with decision time. (C–D) Predictions of the value cut-off heuristic and differences cut-off heuristic. The integrated-evidence model predictions are shown for low cutoffs (cyan lines) and for high cut-offs (magenta lines). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

access to). However, as we show below, we can make model-predictions on how this variable should depend on decision-time, as well as on the range of internal noise given the actual response and its RT for a specific trial. Fig. 2 (upper panels) presents the predicted average integrated evidence of the correct responses, as a function of type of boundary (fixed, collapsing or no-boundary) and level of internal noise (see Figs. S1–4 for replication of Fig. 2 using different sets of parameters and mixed-difficulty condition).

As expected, under low internal-noise, the fixed-boundary model predicts that the average integrated-evidence changes little with the decision-time, while the collapsing-boundary model predicts a drop of the integrated-evidence with decision-time. Under high internal noise, however, the fixed-boundary model predicts an increasing integrated-evidence at the time of decision (Fig. 2A). The reason for this is the omittance of internal noise (which was integrated together with the evidence to determine the stopping time) from the integrated-evidence, as this noise is correlated with the RT. While in shorter trials the internal-noise happened to be facilitatory (i.e., in the direction of the external evidence, leading to faster responses), in longer trials it happened to be inhibitory (against the external-evidence, leading to slower responses). Thus, when omitting the internal-noise, the integrated-evidence underestimates the boundary on short RT and overestimates it on longer RT. For the case of collapsing boundaries with high internal noise, the two processes balance out, so that the integrated evidence is roughly fixed. The boundary models differ in their predictions from the heuristic and timer models (Fig. 2, bottom panel), which predict a faster increase in the integrated-evidence with time.

Finally, we examine the RT predictions of the different models as a function of difficulty (see Fig. 1B). As shown in Fig. 3A, the value-cutoff heuristic predicts that RT decreases with difficulty (this pattern is obtained across different cut-off values, Fig. S5). This is the outcome of statistical facilitation (Raab, 1962): as the mean of the low Gaussian increases, there is a higher chance to sample a value that exceeds the cut-off. A similar prediction also takes place in models that assume integration of absolute (rather than relative) evidence (Teodorescu & Usher, 2013). The counting and integration-to-boundary models have qualitatively similar predictions, which are in the opposite direction from the value-cutoff: RTs are slower in the difficult conditions

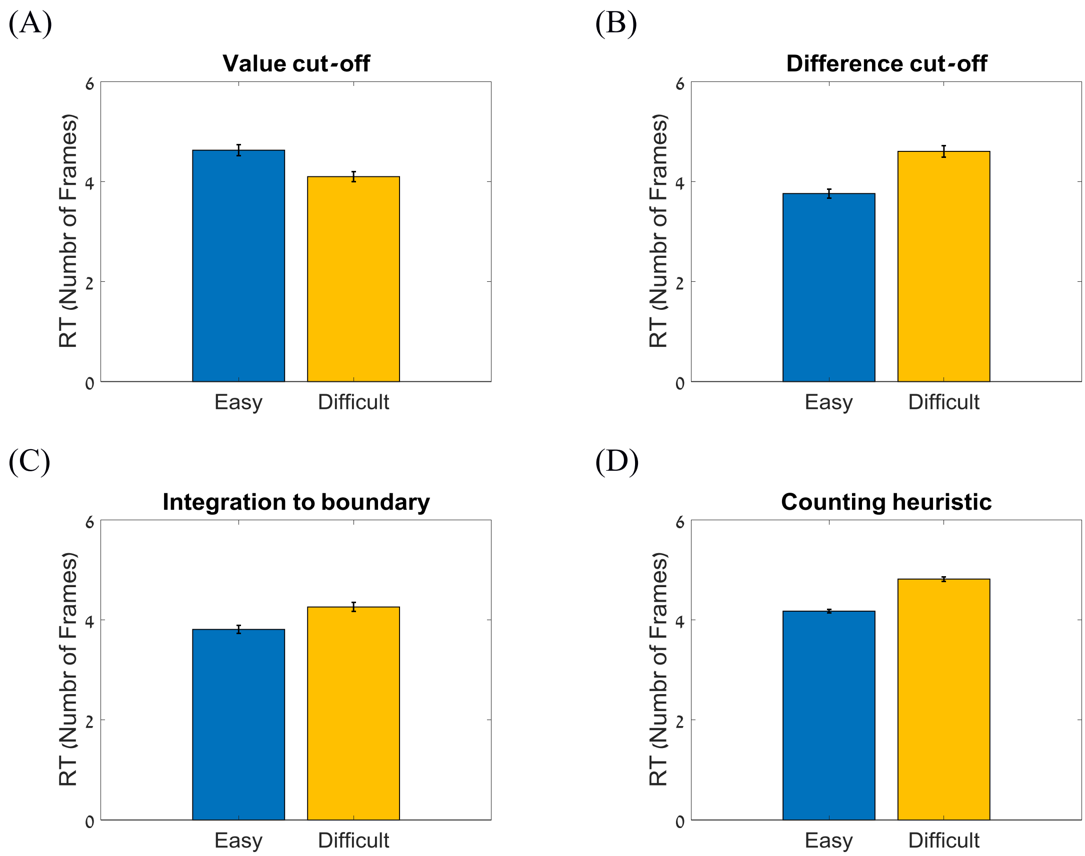

Fig. 3. Mean-RT as a function of task difficulty, predicted by (A) Value cut-off heuristic, (B) Difference cut-off heuristic, (C) Integration to boundary model, and (D) Counting heuristic (absolute criterion). Error bars correspond to 95% confidence intervals.

(Fig. 3B–D). We thus distinguish those using quantitative model-fits in the next section.

3. Experiments

3.1. General experimental methods

Both experiments used the same experimental paradigm (Fig. 1). Two rapidly changing (2 Hz) sequences of numerical values (Glickman et al., 2018, Tsetsos et al., 2012) were presented to the participants, who were asked to indicate which sequence was drawn from a distribution with a higher mean. Critically, we used a free response paradigm, in which the decision terminates the trial (Smith & Vickers, 1989). We used two difficulty levels: easy trials, in which the means of the Gaussians were: $\mu_1 = 52$ vs. $\mu_2 = 44$, $\sigma = 10$, and difficult trials in which the means were: $\mu_1 = 52$ vs. $\mu_2 = 48$, $\sigma = 10$. The two experiments differ in the reward instructions given to the participants. All subjects participated in both experiments in different sessions (separated by about 1 week). See *Suppl. Methods* for additional details.

3.2. Experiment 1

3.2.1. Participants

Twenty-seven undergraduate from Tel-Aviv University (22 females; age: $M = 23$, range 21–28 years) participated in the experiment. The participants received course credit in exchange for taking part in the experiment, as well as a bonus fee ranging from 15 to 25 ILS, which was determined by their task performance. The experiment was approved the ethics committee at TAU.

3.2.2. Model-fit procedure

The models were fitted to the choice and RT-data, based on the actual sequences that were sampled in each trial, using maximum likelihood estimations (see *Suppl. Model Fitting*).

3.2.3. Results

First, we compared the accuracy and decision-times of the easy and difficult trials. As shown in Fig. 4A, the participants were more accurate ($t(26) = 22.07$, $p < .001$) and responded faster ($t(26) = 7.0$, $p < .001$) on the easier trials. These findings are consistent with a competitive (rather than independent) integration of evidence, which predicts slowdown in response time with increased difficulty (Teodorescu & Usher, 2013). The results also speak against a value cut-off model, which also predicts faster responses in the more difficult trials (Fig. 3A). Interestingly, as shown in Fig. 4B (black line), we find that accuracy shows a small drop with RT. This is contrary to what a timer model would predict (increase in accuracy with the number of samples), and also challenging for a fixed boundary model (which predicts a roughly constant accuracy; see below).

Next, we examined how the average integrated evidence at the time of response varies as a function of decision-time. Note that our experimental paradigm allowed us to monitor the rate of evidence accumulation over the course of each trial, and thus we were able to compute the exact integrated evidence at the time of decision. Fig. 4C presents the data of a representative participant, showing an approximately constant value of average integrated evidence across time (see Fig. S6 for the results of all participants). The results at the group level were evaluated using a mixed-effect linear regression, with random intercept and slope for each participant, which also showed time-

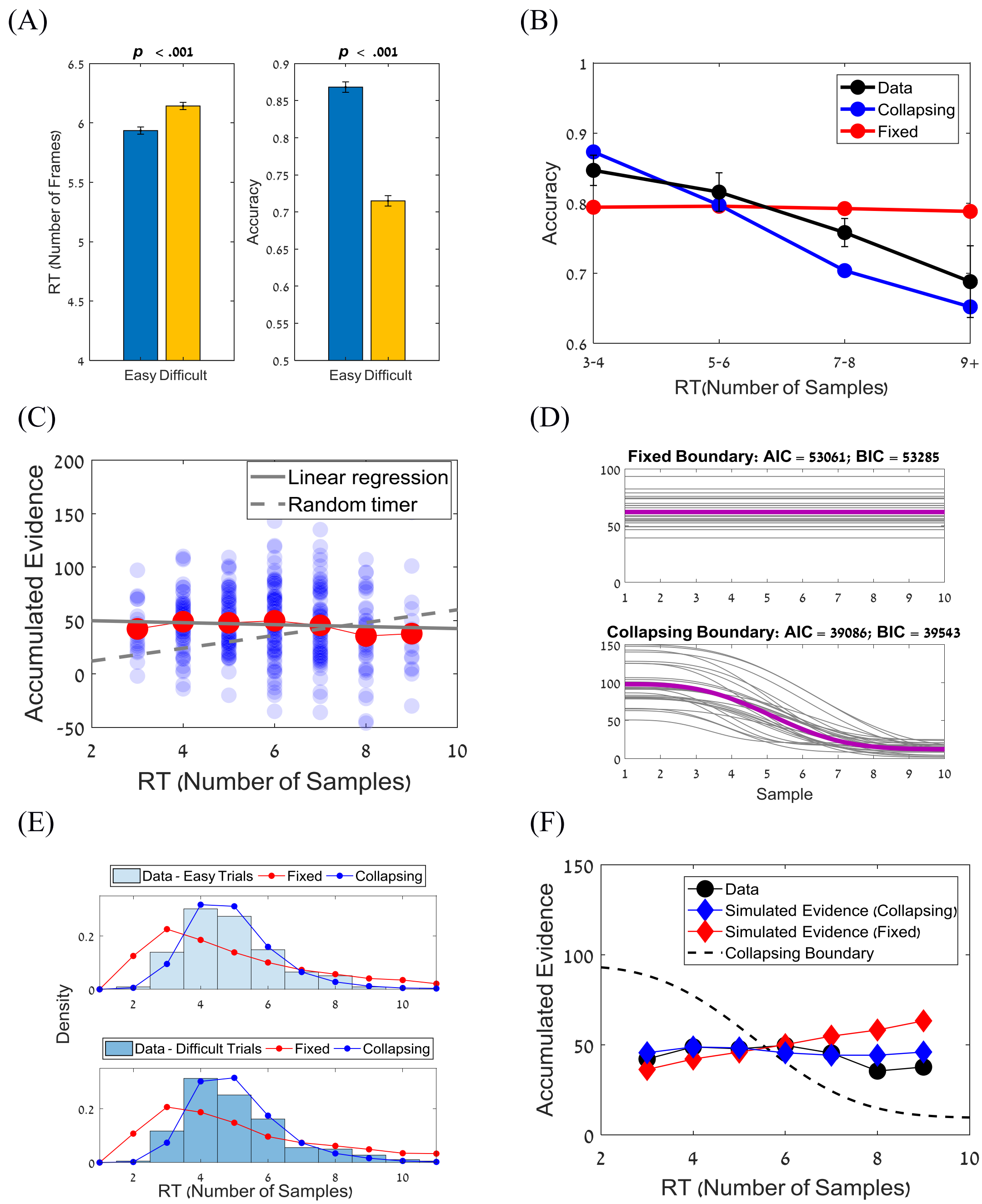

Fig. 4. Results of Exp. 1. (A) The participants were more accurate and faster on easy trials than on difficult ones. (B) Accuracy as a function of response-time. Black line represents the empirical accuracy; Red/Blue lines represent the simulated accuracy of the fixed/collapsing boundary models, respectively. (C) Scatter-plot showing the accumulated evidence in individual trials (blue circles) as a function of decision time of a representative participant; red circles correspond to the average accumulated evidence at each RT, the gray solid line corresponds to linear regression fitted to the data, and the dashed gray line corresponds to the prediction of the timer model. (D) The collapsing-boundary model decisively outperformed the fixed-boundary model; Solid purple lines correspond to the boundaries generated using the group mean parameters, gray lines correspond to the boundaries of the individual participants. (E) Decision-time density distributions of a representative participant as a function of trial difficulty. Red/Blue lines represent simulated decision-time distributions of the diffusion models with fixed/collapsing boundaries, respectively. (F) The match between the actual (black circles) and simulated (blue diamonds for collapsing-boundary and red diamonds for fixed-boundary) integrated evidence of a representative participant. The black dashed line corresponds to the collapsing-boundary generated using the participant best-fitted parameters. Error bars correspond to 95% confidence intervals. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

invariant pattern ($\beta = -0.22, p = .12$; Fig. 5C, blue line). The obtained results are inconsistent with the qualitative predictions of the timer model (Fig. 5C, dashed rising line), but are rather consistent with either: i) fixed-boundary without (or with low) internal noise or with ii) collapsing-boundary with internal noise.

To distinguish between the fixed and collapsing boundary models, we carried out a model comparison, which estimated the likelihood of each model given the subject choice and the samples inspected in each trial. The model comparison provided decisive support to the collapsing-boundary model over the fixed-boundary model, as indicated by lower *AIC* and *BIC* scores for all the participants, as well as by a decisive support at the group level (ΔAIC = 13,975, ΔBIC = 13,742; Fig. 4D, Tables S1–2). Similar results were obtained when we compared the collapsing-boundary model to a fixed-boundary model, which includes

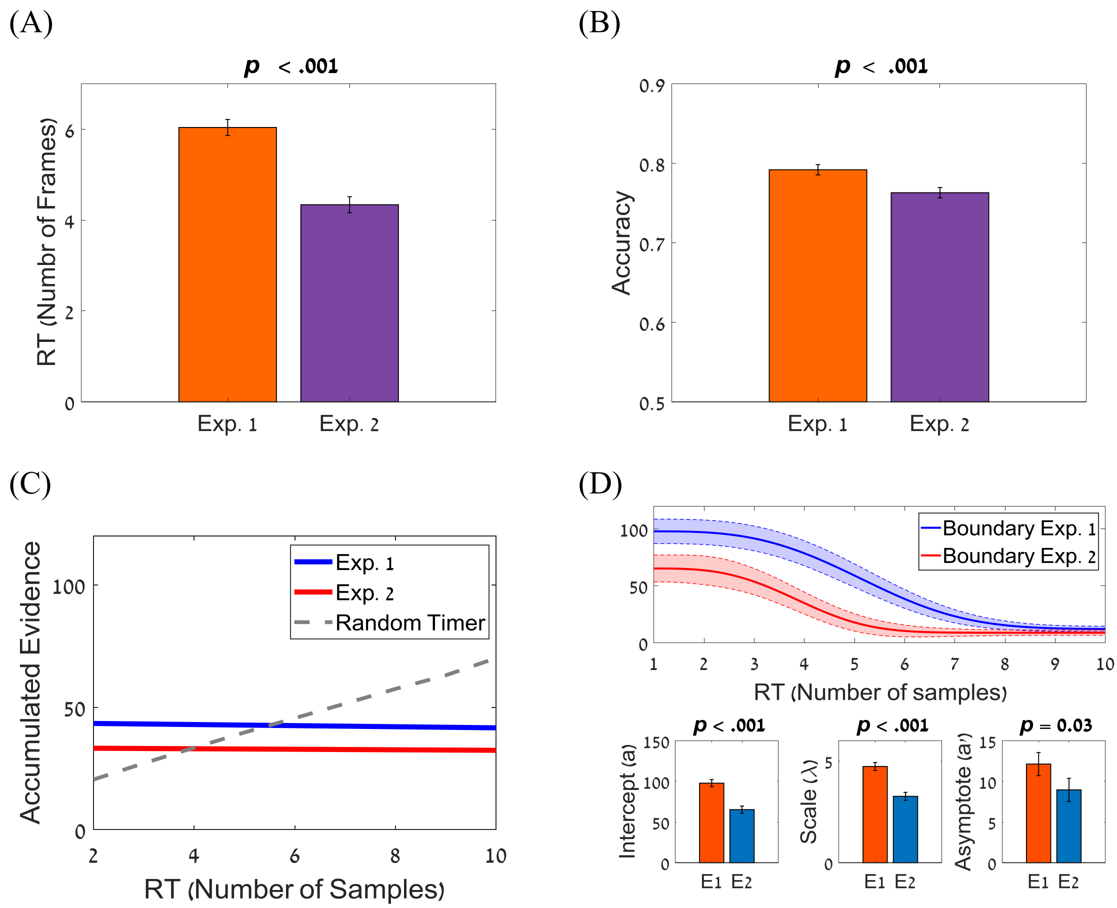

Fig. 5. Results of Exp. 2. (A-B) The accuracy in Exp. 2 was lower than in Exp. 1, and decision-times were higher as a result of the payoff-manipulation. (C) The accumulated evidence in Exp. 2 (red line) were lower than Exp. 1 (blue line), however, both showed time-invariant pattern which is not consistent with the predictions of the timer model (gray dashed line). (D) Upper panel: the mean boundary obtained in Exp. 1 was higher than the one obtained in Exp. 2. Lowe panel: the average intercept, scale and asymptotic parameters obtained in Exp. 2 were lower than those obtained in Exp. 1. Error bars correspond to 95% confidence intervals. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

between trial variability parameters for starting point and drift rate ($\Delta AIC = \Delta BIC = 13{,}933$). We also tested the two best counting heuristic variants (using collapsing boundaries), which provided a much lower fit to the results ($\Delta AIC = \Delta BIC = 3214$ for the absolute criterion and $\Delta AIC = \Delta BIC = 1589$ for the relative criterion; Tables S1–2).

Finally, we examined the accounts provided by the fitted integration to boundary models (fixed/collapsing) to the empirical data. The integration to a collapsing-boundary model accounted for the accuracy as a function of RT (Fig. 4B, blue line) and for the distribution of decision-times as a function of difficulty (Fig. 4E, blue line).[2] The fits are somewhat less successful for the fixed boundary model (red line), which does not account for the reduced accuracy with RT, and overestimates the leading-edge and tail of the RT-distributions. A novel aspect of our computational method is that we fit the choice and RT for each trial given the actual evidence (in that trial). As we show in the Supplement (Figs. S9–10), the simulated choices and RT show an increased correlation with the data, compared to simulations based on randomly sampled sequences, shuffled across trials with the same statistical properties.

In order to reconcile the apparent discrepancy between the constant value of the average accumulated evidence and the concave shape of the boundaries, we simulated the fixed and collapsing boundary models for each participant using his or her best-fitted parameters. The results demonstrate a good match between the simulated (Fig. 4F, blue line) and actual (black line) integrated evidence, suggesting that despite the collapsing boundaries, the integrated-evidence (excluding the noise) remains constant across time (see Fig. S11 for the results of all participants). On the other hand, the fixed boundary model overestimated the accumulated evidence for a high number of samples (red line).

3.3. *Experiment 2*

The aim of the experiment was to replicate the integration-to-boundary results, while also manipulating the response-boundary by introducing a cost for the decision-time.

3.3.1. *Participants*

The same participants as in Exp. 1, did a second session one week later.

3.3.2. *Reward instructions*

The reward decreased with the number of frames inspected. At each trial the reward was 10–'# inspected-frames' points, for each correct response, and 0 points for errors. Points were translated into monetary reward at the end of the experiment.

[2] In all model fits we assumed that the participants' responses were based on the model examined for *all trials*. However, it is possible that in a small minority of trials, the participants made guessing responses (Ratcliff & Tuerlinckx, 2002) or may responded based on information related to previous trials (Braun, Urai, & Donner, 2018). Exclusion of such trials may further improve the model fit of the collapsing boundary model in Fig. 4B, but we prefer to focus here on a standard model comparison without outlier exclusion, or model extension to include sequential dependencies, which require a separate investigation.

3.3.3. Results

We first replicated the results obtained in Exp. 1: the participants responded faster ($t(26) = 3.91$, $p = .006$) and were more accurate ($t(26) = 17.43$, $p < .001$) in the easier compared with the more difficult trials. Next, we examined the effect of our payoff manipulation on accuracy and decision-times. As predicted, the participants responded faster ($t(26) = 9.89$, $p < .001$; Fig. 5A) and were less accurate in Exp. 2 ($t(26) = 4.33$, $p < .001$; Fig. 5B), indicating sensitivity to our payoff manipulation. Then, we compared the integrated evidence at the time of response between the two experiments. The results replicated the time-invariant pattern obtained in Exp. 1 ($\beta = -0.10$, $p = .51$), however, because of the payoff manipulation, the participants used lower levels of integrated evidence (Fig. 5C, red line).

Second, we repeated the model comparison which was conducted in Exp. 1. As in Exp. 1 the collapsing-boundaries model outperformed the fixed-boundaries model ($\Delta AIC = 7973$, $\Delta BIC = 7744$). Importantly, we found that the best-fitted parameters obtained in Exp. 1 were highly correlated with those obtained in Exp. 2, indicating consistent integration dynamic (Fig. S12).

Critically, the comparison of the two sets of parameters revealed significant differences in the boundary parameters (Fig. 5D). In order to speed up their RT, the participants used lower intercepts ($t(26) = 8.61$, $p < .001$), collapsed their boundaries faster ($t(26) = 8.16$, $p < .001$), and their boundary collapsed to lower asymptotic ($t(26) = 2.28$, $p = .03$; Fig. 5D).

4. Discussion

In this study we provided evidence for integration-to-a-collapsing-boundary in a choice task with rapid numerical values presented until response, and for adaptation of the boundaries to reward contingency. The experimental results (Exp. 1) showed that task-difficulty reduced accuracy and increased RT as predicted by models of relative evidence-integration (Teodorescu & Usher, 2013), and in opposition to a value-cutoff heuristic that make the opposite prediction (Fig. 3). Additionally, we found that the average integrated-evidence at the time of decision is roughly time-invariant, in contradiction to a random-timer decision-termination mechanism and the cutoff heuristics (Fig. 2C–D), and that the average integrated-evidence is lower in response to the time-costing reward contingency (Exp. 2; Fig. 5C). Using computational modeling, we were able to confirm these results (and rule out counting models) and identify the type of boundary that participants rely on. Model comparison provided strong support for a collapsing-boundary model, whose boundary is reduced and collapses faster under the time-cost reward contingency (Fig. 5D). Moreover, this model (but not the fixed-boundary) provides an excellent account for the decision-time distributions (Fig. 4E, S7–8) and for the speed-accuracy function (Fig. 4B). We note that in order to account for the time-invariant accumulated evidence at response (Fig. 4C), the internal noise in the collapsing-boundary model is relatively large (about a factor of 1.3 higher than the external sampling noise).

The exact nature of the decision boundaries has been subject to considerable debate, with empirical evidence producing mixed results. While a number of studies found evidence supporting time invariant boundaries (Hawkins et al., 2015; Voskuilen, Ratcliff, & Smith, 2016), others provided support for a time-variant integration policy (Drugowitsch et al., 2012; Khodadadi, Fakhari, & Busemeyer, 2017; Palestro et al., 2018). There are a number of sources for these discrepancies, ranging from the time-scale of evidence integration Malhotra et al., 2017) to optimality considerations. In particular, Malhotra et al. (2017) have shown that it is optimal to collapse the decision boundary when trials of mixed difficulty are randomly presented and that people can adapt the boundary policy to this experimental contingency in an expanded judgment task. As our task involved mixed difficulty trials and the time scale of integration was relatively long, our results are consistent with the idea that, at least on this time scale, people can adopt a roughly optimal policy. Note that while Exp. 1 posed no explicit time-cost, it is likely that implicit costs exists (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006; Edwards, 1965). Moreover the reduction in the boundary parameters obtained when a cost for time was introduced (Exp. 2), further demonstrates the adaptive character of the decision mechanism.

Unlike previous studies, which examined evidence integration with perceptual-evidence, here we used rapid numerical sequences (Glickman et al., 2018; Tsetsos et al., 2012). Such sequences are relevant to value based choices, such as choosing between stocks based on sequences of returns. We reasoned that within 500 ms, symbolic number-values can be associated with representations of magnitudes (Brezis, Bronfman, & Usher, 2015; Dehaene, 2011) and therefore can provide the input to an optimal choice mechanism based on integration-to-boundary. The broad tuning of such magnitude representations is consistent with the relatively large internal noise in the collapsing-boundary model (Brezis, Bronfman, Jacoby, Lavidor, & Usher, 2016). As shown in Fig. S13, the average accuracy achieved by our subjects, who enhanced signal-to-noise ratio by integrating to a criterion, is 87% (group-level) of that achievable by an ideal observer (that integrates the differences without internal noise). Obviously there are non-optimal strategies to carry-out our task, such as cutoff or counting heuristics or a random-time integration policy. The results are thus remarkable, as they refute these alternative strategies, providing strong support for integration-to-boundary even with numerical stimuli. Future research will need to examine how the boundary parameters depend on type of evidence (perceptual vs. numerical) or alternatively on presentation-rate of the alternatives.

Acknowledgments

This research was supported by grants to Marius Usher from the United States-Israel Binational Science Foundation (CNCRS = 2014612), and from the Israel Science Foundation (1413/17). We wish to thank Zohar Bronfman and Konstantinos Tsetsos for helpful discussions and Rani Moran for comments on our manuscript and very helpful suggestions on model fitting.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2019.104022.

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*(4), 700–765.

Braun, A., Urai, A. E., & Donner, T. H. (2018). Adaptive history biases result from confidence-weighted accumulation of past choices. *Journal of Neuroscience, 38*(10), 2418–2429.

Brezis, N., Bronfman, Z. Z., Jacoby, N., Lavidor, M., & Usher, M. (2016). Transcranial direct current stimulation over the parietal cortex improves approximate numerical averaging. *Journal of Cognitive Neuroscience, 28*(11), 1700–1713.

Brezis, N., Bronfman, Z. Z., & Usher, M. (2015). Adaptive spontaneous transitions between two mechanisms of numerical averaging. *Scientific Reports, 5*, 10415.

Brusovansky, M., Vanunu, Y., & Usher, M. (2017). Why We Should Quit While We're Ahead: When Do Averages Matter More Than Sums?.

Dehaene, S. (2011). *The number sense: How the mind creates mathematics.* OUP USA.

Dehaene, S., Molko, N., Cohen, L., & Wilson, A. J. (2004). Arithmetic and the brain. *Current Opinion in Neurobiology, 14*(2), 218–224.

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

Edwards, W. (1965). Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing. *Journal of Mathematical Psychology, 2*, 312–329.

Glickman, M., Tsetsos, K., & Usher, M. (2018). Attentional selection mediates framing and risk-bias effects. *Psychological Science, 29*(12), 2010–2019.

Gold, J. I., & Shadlen, M. N. (2002). Banburismus and the brain: Decoding the relationship between sensory stimuli, decisions, and reward. *Neuron, 36*(2), 299–308.

Hawkins, G. E., Forstmann, B. U., Wagenmakers, E. J., Ratcliff, R., & Brown, S. D. (2015).

Revisiting the evidence for collapsing boundaries and urgency signals in perceptual decision-making. *Journal of Neuroscience, 35*(6), 2476–2484.

Khodadadi, A., Fakhari, P., & Busemeyer, J. R. (2017). Learning to allocate limited time to decisions with different expected outcomes. *Cognitive psychology, 95*, 17–49.

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience, 13*(10), 1292–1298.

Malhotra, G., Leslie, D. S., Ludwig, C. J., & Bogacz, R. (2017). Overcoming indecision by changing the decision boundary. *Journal of Experimental Psychology: General, 146*(6), 776–805.

Mazurek, M. E., Roitman, J. D., Ditterich, J., & Shadlen, M. N. (2003). A role for neural integrators in perceptual decision making. *Cerebral cortex, 13*(11), 1257–1269.

Palestro, J. J., Weichart, E., Sederberg, P. B., & Turner, B. M. (2018). Some task demands induce collapsing bounds: Evidence from a behavioral analysis. *Psychonomic Bulletin & Review,* 1–24.

Raab, D. H. (1962). Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences, 24*(5), 574–590.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59–108.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20*(4), 873–922.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science, 9*(5), 347–356.

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.

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review, 9*(3), 438–481.

Smith, P. L., Ratcliff, R., & Sewell, D. K. (2014). Modeling perceptual discrimination in dynamic noise: Time-changed diffusion and release from inhibition. *Journal of Mathematical Psychology, 59*, 95–113.

Smith, P. L., & Vickers, D. (1989). Modeling evidence accumulation with partial loss in expanded judgment. *Journal of Experimental Psychology: Human Perception and Performance, 15*(4), 797.

Spitzer, B., Waschke, L., & Summerfield, C. (2017). Selective overweighting of larger magnitudes during noisy numerical comparison. *Nature Human Behaviour, 1*(8), 0145.

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

Tsetsos, K., Chater, N., & Usher, M. (2012). Salience driven value integration explains decision biases and preference reversal. *Proceedings of the National Academy of Sciences, 109*(24), 9659–9664.

Vanunu, Y., Pachur, T., & Usher, M. (2018). Constructing preference from sequential samples: The impact of evaluation format on risk attitudes. *Decision, 6*(3), 223–236.

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.

Wagenmakers, E. J., Van Der Maas, H. L., & Grasman, R. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review, 14*(1), 3–22.

Wald, A. (1947). Foundations of a general theory of sequential decision functions. *Econometrica, Journal of the Econometric Society,* 279–313.

Zeigenfuse, M. D., Pleskac, T. J., & Liu, T. (2014). Rapid decisions from experience. *Cognition, 131*(2), 181–194.