Visual fixations determine the retrieval rate from memory in decision making

Takao Noguchi, Bradley C. Love

Department of Experimental Psychology, University College London

Author Note

Word count: 2196.

Takao Noguchi, Department of Experimental Psychology, University College London; Bradley C. Love, Department of Experimental Psychology, University College London;

Takao Noguchi and Bradley C. Love were supported by the Leverhulme Trust grant RPG–2014–075.

Correspondence concerning this manuscript should be addressed to Takao Noguchi, Department of Experimental Psychology, University College London, WC1E 6BT, UK. Tel: +44 (0)20 3108 55205. E-mail: t.noguchi@ucl.ac.uk

The data can be downloaded at ***.

Abstract

Across a number of domains, drift-diffusion models (DDMs) successfully account for the accuracy and speed of decisions. These models are typically applied to data from perceptual judgments in which information is externally sampled from a stimulus until the accumulated evidence exceeds a boundary, leading to a decision. Here, we consider how DDMs apply to categorization decisions that involve both external sampling of the stimulus and internal sampling of memory. Unlike perceptual judgments, evidence for categorization decisions resides in memory. The stimulus does not provide the evidence, but instead serves as a retrieval cue for relevant memories that provide evidence. By monitoring eye movements, we test the hypothesis that the rate of evidence accumulation should change within a categorization decision as the representation of the stimulus is constructed. At each moment, the previously fixated aspects of stimulus should jointly direct internal sampling from memory. In support of this hypothesis, we find that the order of visual fixations determines the speed and accuracy of categorization decisions in the predicted manner — the earlier within a trial that category atypical information is encoded, the slower and less accurate the decision will be.

Keywords: decision making; evidence accumulation; sequential sampling; categorization; eye movements

Visual fixations determine the retrieval rate from memory in decision making

One unifying idea in decision making research is that people sequentially process information when making a decision, whether that information is a visual input, such as when determining whether a car is approaching, or a verbal input such as when evaluating whether a company is a good investment. The drift-diffusion model (DDM) has proven successful in capturing the time course and accuracy of such decisions (e.g., Bogacz, Wagenmakers, Forstmann, & Nieuwenhuis, 2010; Ratcliff & Smith, 2004). DDMs hold that information is processed sequentially such that evidence is accumulated until an internal boundary is crossed. DDMs are useful in understanding diverse phenomena ranging from perceptual decision making in non-human primates (e.g., Gold & Shadlen, 2007; Kable & Glimcher, 2009) to value-based decisions in humans (e.g., Krajbich, Armel, & Rangel, 2010; Krajbich & Rangel, 2011).

Closely related to our contribution, DDMs can be used to understand how people sample their own memories when making a decision. In many situations, people may sequentially sample a small set of related examples from memory to guide a decision (e.g., Giguére & Love, 2013; Tversky & Kahneman, 1974; Vlaev, Chater, Stewart, & Brown, 2011). For example, radiologist may view a mammogram and be reminded of past patients and these sampled memories may guide diagnosis (Hornsby & Love, 2014). More prosaically, a dark cloud may elicit memories of heavy rains, leading people to pack an umbrella instead of sunglasses. The process of sequential sampling from memory, evidence accumulation, and decision has been formalized in the exemplar-based random walk (EBRW) model (Nosofsky & Palmeri, 1997), which marries a model of categorization (and concordant memory representations) with a DDM.

What one samples from memory to guide a decision should depend on which aspects of the situation are encoded. For example, imagine you are attending a rock concert and are trying to decide whether the person standing in front of you is male or female. After noticing the person has long hair, you sample mostly female examples from memory.

However, the concert goer turns and you notice a full beard, which cues sampling of mostly male examples from memory, driving you to correctly categorize the concert goer as male. In this case, the rate at which evidence is accumulated toward the competing alternatives (male vs. female) changes as a function of which dimensions (hair length, presence of beard, etc.) of the stimulus are encoded. As new information is encoded, the similarity relations to items in memory are altered, which changes the probability that a memory will be sampled and serve as evidence to shape the decision.

This dynamic construction of stimulus representation contrasts with typical DDMs in which the average rate (referred to as the drift rate) that accumulated evidence grows (toward the correct decision boundary) is constant within a trial (i.e., for a particular decision). In contrast, as the previous example makes clear, the rate and direction of evidence accumulation may change as a function of which aspects of the situation are encoded.

In this contribution, we consider this linkage between external sampling of the stimulus and internal sampling of memory by monitoring eye movements while people make categorization decisions. As people attend to various dimensions of a stimulus, the accumulation rate may dynamically change in predictable ways. Indeed, previous studies find that people under time pressure are likely to evaluate a stimulus not as whole but piece-by-piece as dimensions are attended (e.g., Cohen & Nosofsky, 2003; Lamberts, 1995, 2000). These findings indicate that accumulation rate may be based on a subset of dimensions first with more dimensions gradually being incorporated.

The hypothesis we forward, that people's sequential sampling of the external visual stimulus directs their internal sampling of memory providing evidence for the decision, makes a number of testable behavioral predictions. To illustrate, suppose a stimulus has *consistent* and *inconsistent* values on stimulus dimensions with respect to its category membership. For example, for the dimension of hair length, males can have long hair and females short hair, but long hair is more consistent with the female category. A consistent value on a dimension will tend to drive accumulation toward the correct boundary, whereas an inconsistent value will drive accumulation toward the incorrect boundary. These basic effects of inconsistent and consistent dimension values should hold for all decisions, irrespective of whether information directly from the visual stream or sampled memories serve as evidence.

However, the predictions are more subtle and interesting for the case we focus on in which decisions are made by sampling memory. When gathering evidence from memory, the order in which perceptual information is externally sampled (e.g., eye movements) is critical. The reason is that the attended aspects of the stimulus are predicted to continuously serve as a cue throughout the decision such that the earlier in a decision an inconsistent piece of information is sampled, the more the decision maker should be misled. In effect, the longer the misleading information is known, the more opportunities there should be to sample misleading memories before a decision bound is crossed.

The predicted effects are illustrated in Figure 1 (a). The two paths in this figure represents different attention orders. When an inconsistent value is attended first (blue line), evidence tends to accumulate towards the incorrect boundary. To make a correct decision, this incorrect accumulation has to be reversed, which increases the number of steps (i.e., time, memory sampling operations) required to reach the correct boundary. Therefore, attention orders influence the speed to make a correct decision. In contrast, when stimulus representation is not constructed and accumulation only depends on current attention, attention orders do not influence the speed (see Figure 1 (b)).

Further with the dynamic construction, it is possible that the incorrect decision boundary is crossed, resulting in an error, before the effect of the consistent information can reverse the initial misleading information. This prediction on accuracy makes contrast to the static representation, where accumulation starts only after stimulus is completely represented (Figure 1 (c)). In addition, the dynamic construction predicts that when the first two initial eye movements are to dimensions with values consistent with the category,

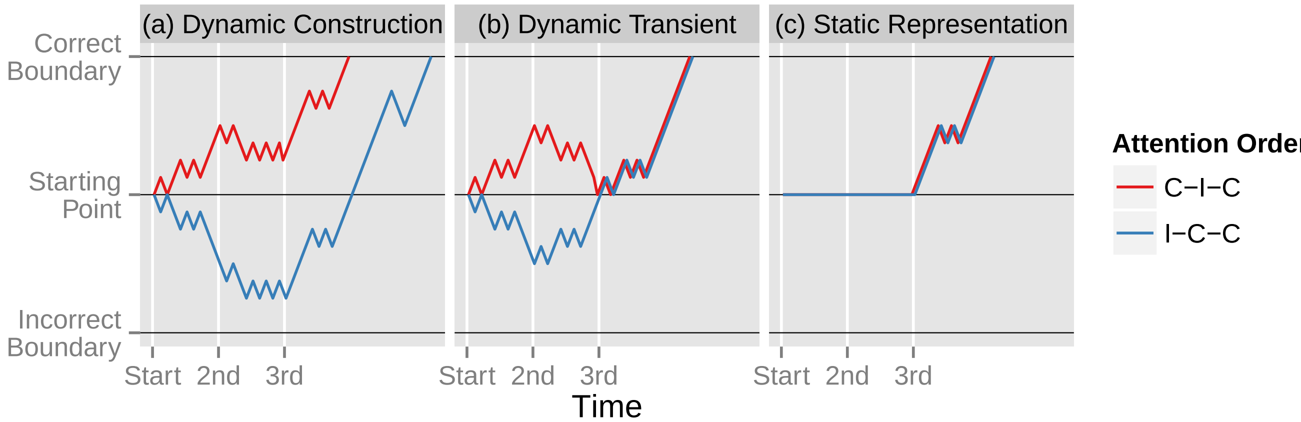

Figure 1. The predicted effects of attention order on speed and accuracy of decision for three competing constructions of stimulus representation. Along the horizontal axis, “2nd” indicates the time point where the second dimension is attended and incorporated into an accumulation rate. Similarly, “3rd” indicates the time point where the third dimension is attended and incorporated. “C” in the legend indicates attention to a consistent value, and “I” indicates attention to an inconsistent value. Therefore, “I-C-C” indicates attention to an inconsistent value first, followed by attention to consistent values. (a) The dynamic construction predicts that previously and currently attended values serve as the retrieval cue, leading to attention order influencing the speed and accuracy of decisions. In contrast, (b) the dynamic transient predicts that only a currently attended value serves as the retrieval cue, leading to attention order influencing only the accuracy of decisions. Also, (c) the static representation predicts that the retrieval starts only after all the values are attended, and hence that the speed and accuracy of decisions are independent of attention order. All three constructions hold that the stimulus serves as the retrieval cue, but the views differ from each other on how the stimulus guides retrieval.

participants may reach the decision boundary and respond prior to making a third eye movement to the remaining stimulus dimension (cf. Blair, Watson, & Meier, 2009; Blair, Watson, Walshe, & Maj, 2009).

We conducted a study to test these predictions in a categorization task in which stimuli contained three dimensions (i.e., informative locations) and eye movements were monitored. To foreshadow, the main predictions regarding the order of stimulus sampling (i.e., eye movements) and behavioral performance held, supporting the notion that people are sampling memory dynamically when making categorization decisions.

Methods

Participants

Seventy two participants were recruited from the participant pool at the University of Warwick. One participant could not complete the experiment due to failure in tracking their eye-movements, leaving 71 (33 male and 38 female) participants. Their age ranged from 18 to 33 with a mean of 20.99.

Stimuli and Design

The stimuli we used in the study are taken from Blair, Watson, Walshe, and Maj (2009). An example stimulus is illustrated in the left panel in Figure 2. Each stimulus contains three dimensions, which manifest one of two possible values. The right panel in Figure 2 illustrates the category structure used, which consisted of eight items equally divided between Categories A and B. In this family-resemblance structure, each dimension has a characteristic (i.e., consistent) value for each category. For example, the second member of Category A (see Figure 2) has an inconsistent value on the top dimension — although this item belongs to Category A, the value of its top dimension is more common of Category B. Thus, when only the top dimension is considered, this Category A item is more similar to Category B items than Category A items. The category structure used is equivalent to Type IV problem (Shepard, Hovland, & Jenkins, 1961) where each dimension is an imperfect predictor of category membership. Each category member displays two or three dimension values consistent with its category. The two possible values of each

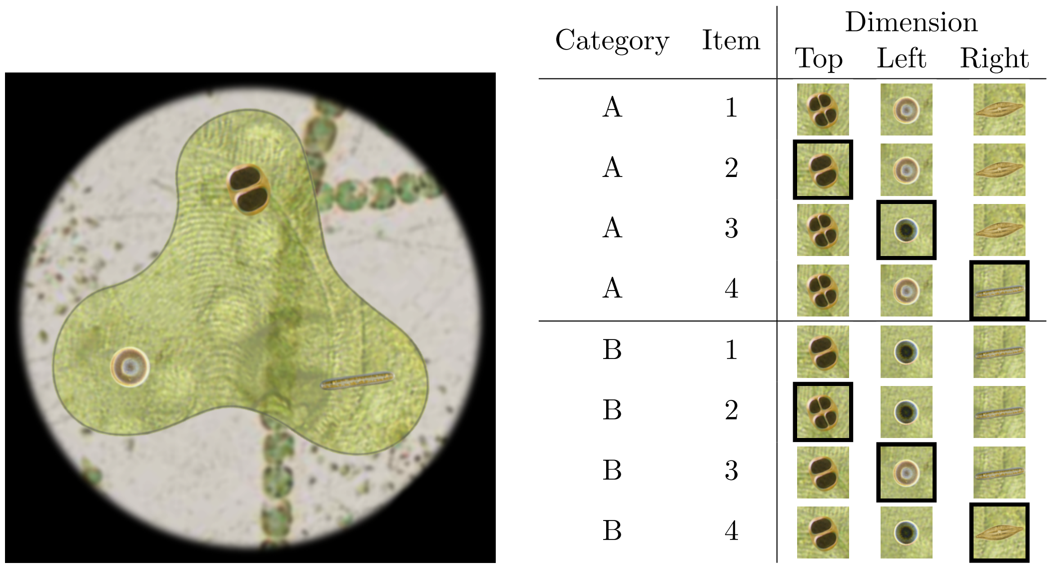

Figure 2. An example stimulus (left panel) and the category structure used (right panel). In the left panel, the three informative stimulus dimensions reside at the top, left, and right locations within the stimulus. In the right panel, black squares (not shown to participants) indicate dimensions that have inconsistent values for the category such that the value is more common to members of the opposing category. This category structure is equivalent to Type IV problem (Shepard, Hovland, & Jenkins, 1961).

dimension were held constant across trials within a participant (e.g., the top location always displayed one of two contrasting values for a given participant) and these location-value pairings were randomly assigned for each participant.

Procedure

Participants learned about the categories through trial-and-error learning: stimulus, then response, then corrective feedback. Participants completed 36 trial blocks. In each block, each of the eight stimuli were presented in a random order.

Each trial began with a fixation cross appearing at the center of the screen. After the participant fixated the cross for 500ms, the stimulus was displayed. The participant judged whether the stimulus belonged to Category A or B by pressing Z key or M key, respectively.

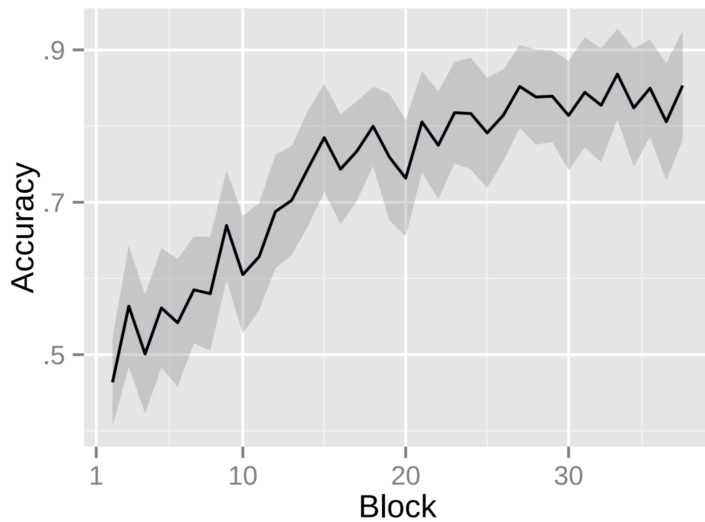

Figure 3. Categorization performance over the 36 blocks. Accuracy is computed for each block for each participant and mean-averaged across participants. The shaded area represents the 95% confidence interval of this mean.

When a participant indicated a decision, feedback ("Correct" in blue font or "Incorrect" in red font) was immediately presented. Instructions emphasized both speed and accuracy.

Throughout the experiment, participant's eye-movements were recorded at 500Hz using an EyeLink 1000 (SR research). The eye-tracker was placed under the 19 inch monitor, and the distance between participant's eye and the eye-tracker was 50–55cm. We did not hide values on stimulus dimensions outside visual fixations, making the measure of attention potentially noisy. However, we aimed to reduce noise in eye-tracking by calibrating the eye-tracker frequently: the eye-tracker was calibrated before the experiment and also after every 10 trials during the experiment. For the eye-movement recordings, we defined non-overlapping regions of interest to identify which dimension (see Figure 2) the participant attended.

Results

First, we assessed whether participants learned to make correct decisions in the experiment. The aggregated learning curve is plotted in Figure 3. For each participant, we identified the first two consecutive blocks in which all the items were categorized correctly. Thirty five out of the 71 participants failed to meet this criterion. For the remaining 36 participants, the first of the two consecutive blocks ranged from Block 3 to 33, with a median of 13. Only the trials in these blocks and subsequent blocks are included in the following analyses.

All the statistics we report below are based on χ^2 tests on fit of maximal mixed-effect regressions, which allow each participant and each block to have varying effects of attention order. The estimated hyper-parameters are reported as β.

First, we consider a prediction by the dynamic construction concerning the conditions under which participants should respond prior to viewing all three stimulus dimensions. As predicted, participants were more likely (.28 vs. .15), to make a decision after only viewing two of three locations when both dimensions had consistent information than when one of the first two dimensions contained inconsistent information, $\chi^2(1) = 21.73$, $\beta = 1.05$, $p < .001$.

Our next analyses focus on eye-movement patterns in which all three dimensions are fixated prior to a decision. First, accuracy of decisions is plotted in the left panel in Figure 4. As predicted, decisions tend to be more accurate the later within a trial an inconsistent value is attended, $\chi^2(2) = 26.53$, $p < .001$. The key prediction for I-C-C vs. C-I-C (see Figure 1) held such that sampling the inconsistent value on the first fixation leads to lower performance than sampling the inconsistent value on the second fixation, $\chi^2(1) = 39.36$, $\beta = 1.22$, $p < .001$.

The right panel in Figure 4 illustrates speed of correct decisions. The distribution of response time is positively skewed, and thus we examined logarithm of response time. As predicted, correct decisions are slower when an inconsistent value is attended earlier within

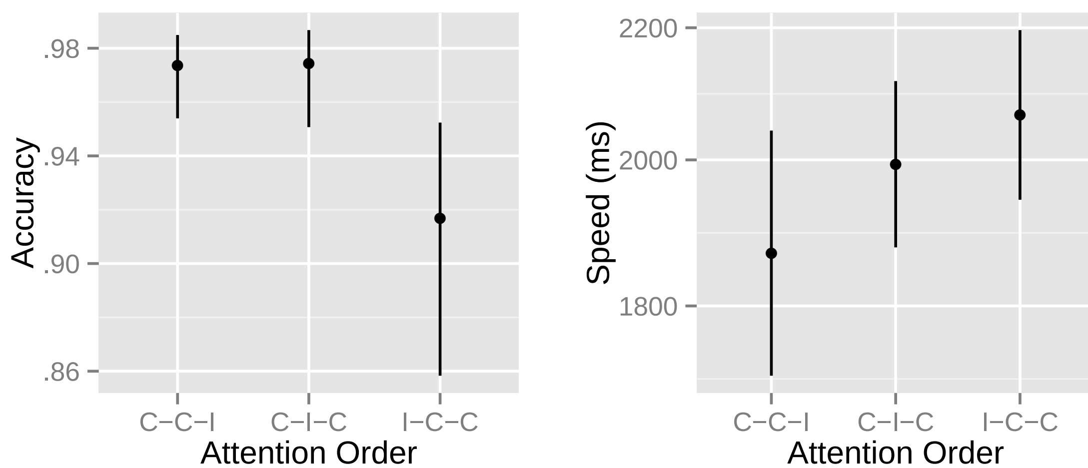

Figure 4. Accuracy (left panel) and speed (right panel) of decision. Along the horizontal axis, "C" indicates a consistent dimension value, and "I" indicates an inconsistent dimension value. Therefore, "I-C-C" indicates attention to an inconsistent piece of information is attended first, followed by two consistent dimension values. Error bar represents 95% confidence interval, estimated with the mixed-effect regressions.

a trial, $\chi^2(2) = 10.05$, $p < .01$. The key prediction for I-C-C vs. C-I-C also held such that sampling the inconsistent value on the first fixation leads to slower performance, $\chi^2(1) = 4.06$, $\beta = 0.04$, $p = .04$.

Discussion

The evidence accumulation underlies a wide range of decisions. Our results provide an insight into the nature of evidence accumulation. In particular, the results show the influences of attention order on speed and accuracy of decisions. Since evidence is accumulated as dimensions are attended, attention to an inconsistent value at first leads to accumulation towards an incorrect boundary. As a result, it takes longer to make a correct decision and it becomes less likely to make a correct decision. These results show that evidence accumulation is a dynamic process with accumulation rate changing as more dimensions are attended.

This dynamicity is due to the linkage between the external sampling of information and the internal sampling of memory. In this linkage, externally sampled information guides what examples are internally more likely sampled. In the Introduction, we discussed the example of deciding whether a person with long hair is male or female. In this example, male examples are are more likely to be sampled internally (i.e., retrieved from memory) when the person's beard has been externally sampled (i.e., visually fixated).

This hypothesized involvement of internal sampling in our study highlights a potential distinction between cognitive and perceptual decisions. According to typical DDMs, for perceptual decisions, evidence accumulation depends on the momentary external sampling of information. Here, accumulation depend not on what *has been* externally sampled but on what *is* being externally sampled (e.g., Gold & Shadlen, 2001). In perceptual decision making, misleading information drives evidence accumulation toward the incorrect boundary, whether or not more representative information has been previously sampled. This independence follows from external, rather than internal sampling, providing evidence for the decision. Our results suggest a different dynamic for cognitive decisions made by sampling from memory. In this case, a representation of the stimulus is constructed incrementally from visual fixations and the current state of this representation determines the rate at which evidence is retrieved from memory.

In summary, our results show how people's external sampling of information is linked to their internal sampling of memory. As examples sampled from memory depend on which aspect of the stimulus has been externally sampled, evidence accumulation is a dynamic process, as shown in the behavioral influences of attention order on cognitive decisions.

References

Blair, M. R., Watson, M. R., & Meier, K. M. (2009). Errors, efficiency, and the interplay between attention and category learning. *Cognition*, *112*, 330–336.

Blair, M. R., Watson, M. R., Walshe, R. C., & Maj, F. (2009). Extremely selective attention: eye-tracking studies of dynamic attention allocation to stimulus features in categorization. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *35*, 1196–1206.

Bogacz, R., Wagenmakers, E.-J., Forstmann, B. U., & Nieuwenhuis, S. (2010). The neural basis of the speed-accuracy tradeoff. *Trends in Neuroscience*, *33*, 10–16.

Cohen, A. L. & Nosofsky, R. M. (2003). An extention of the exemplar-based random-walk model to separable-dimension stimuli. *Journal of Mathematical Psychology*, *47*, 150–165.

Giguére, G. & Love, B. C. (2013). Limits in decision making arise from limits in memory retrieval. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 7613–7618.

Gold, J. I. & Shadlen, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences*, *5*, 10–16.

Gold, J. I. & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, *30*, 535–574.

Hornsby, A. N. & Love, B. C. (2014). Improved classification of mammograms following idealized training. *Journal of Applied Research in Memory and Cognition*, *3*, 72–76.

Kable, J. W. & Glimcher, P. W. (2009). The neurobiology of decision: consensus and controversy. *Neuron*, *63*, 733–745.

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

Krajbich, I. & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings* *of the National Academy of Sciences of the United States of America*, *108*, 13852–13857.

Lamberts, K. (1995). Categorization under time pressure. *Journal of Experimental Psychology: General*, *124*, 161–180.

Lamberts, K. (2000). Information-accumulation theory of speeded categorization. *Psychological Review*, *107*, 227–260.

Nosofsky, R. M. & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, *104*, 266–300.

Ratcliff, R. & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, *111*, 333–367.

Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs*, *75*(13, Whole No. 517).

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science*, *185*, 1124–1131.

Vlaev, I., Chater, N., Stewart, N., & Brown, G. D. A. (2011). Does the brain calculate value? *Trends in Cognitive Sciences*, *15*, 546–554.