**Thresholded Source Memory in Continuous Report: A Diffusion Model Account**

Jason Zhou a, Adam F. Osth a, Simon D. Lilburn a, and Philip L. Smitha

a Melbourne School of Psychological Sciences, The University of Melbourne

Corresponding Author:

Jason Zhou

Melbourne School of Psychological Sciences

The University of Melbourne

Parkville, VIC 3052, AUSTRALIA

jasonz1@student.unimelb.edu.au

Tel: +61 439 603 140

Declaration of Interest: none

Funding sources:

Abstract

**Insert Abstract**

*Keywords:* **Insert Keywords**

Episodic memory, unlike semantic or procedural memory, is memory of a particular event or occasion. A key part of episodic memory is the context in which the event in question occurred. This contextual component of episodic memory is known as *source memory*, and its importance to everyday memory functioning is easily illustrated. For instance, when identifying the suspect of a crime, a memory within the context of a crime scene will have very different implications than a memory of a chance encounter. To successfully complete such an identification task, the witness needs to: 1) encode information about the suspect and the context within his or her memory, 2) bind this information into an overall representation, 3) recognize the suspect at a later time, 4) retrieve the context in which the event occurred, and 5) make a response on the basis of the retrieved representation.

Characterizing each of these components and how they interact is an important question in memory research, and a variety of different models have been proposed to address it. Source memory research has been largely concerned with how contexual cues are used to retrieve the contents of memory and how retrieval varies as a function of the strength of the association between the cue and memory. One specific question that arises about this process is whether the strength of a retrieved memory varies continuously or must exceed a threshold value in order to be retrieved. Several classes of models exist in the literature that formalize these ideas about memory retrieval. Continuous models of source memory predict that retrieved information may be inaccurate but not absent, allowing for a gradual decline in the quality of information retrieved (Banks, 2000; Mickes, Wais & Wixted, 2009). In contrast, threshold or discrete-state models holds that retrieval fails discretely, and so performance is made up of either precise responses, or guesses when the memory is subthreshold (Batchelder & Riefer, 1990; Klauer & Kellen, 2010). A third class of models can be regarded as hybrids of continuous and threshold models , and are known as dual-process models.

Our aim in this article is to compare continuous, threshold, and hybrid models of source memory using a continuous outcome decision task. We present a novel method of analysis of the results of this task using the circular diffusion model of Smith (2016), Our analysis allows us to distinguish the contributions of memory processes and retrieval processes to source memory performance in a precise way.

## Dual-Process Models in Recognition Memory

The debate between continuous and discrete models of source memory is situated in the wider recognition memory literature, which has proposed a range of models to account for memory performance across experimental paradigms. Although source memory is distinct from recognition memory, the two types of memory are often tested simultaneously, and models have been developed to account for both types of tasks (Starns, Hicks, Brown & Martin, 2008). Dual-process models are a class of models which take this integrative approach, in which different kinds of retrieval mechanism support different kinds of memory (Bowers & Schacter, 1990).An influential example is the dual-process model of Yonelinas (1994), in which episodic memory is assumed to involve a mixture of two processes: a fast familiarity-based process and a slower recollection process. Familiarity, defined in this framework, is a quick judgment about whether or not an item has been encountered before based on the strength of its representation in memory (Yonelinas, 1994). Recollection is defined as a slower search process for qualitative information about the item (Atkinson & Juola, 1974). Recollection is assumed to be a thresholded retrieval process while familiarity is a continuous signal-detection process (Yonelinas, 1999). The basis for this assumption is the fact that in recollection people either succeed or fail to retrieve information. While the amount of information above the threshold may vary, Yonelinas (1999) argues that there is a threshold below which there is no information retrieved.

Within a dual-process framework, performance on an item recognition task is assumed to comprise a mixture of both familiarity and recollection processes, in which targets that were studied will be more familiar than lures which were not (Yonelinas, 1994). If successful, recollection enhances recognition by providing details of the study event. Critically, both targets and lures have some level of familiarity, and differ continuously in the degree to which they are familiar, but only targets can be recollected because they are associated with a study event. In source memory tasks, however, familiarity cannot be used to determine the source of items as they have all been studied and are all equally familiar. Consequently, Yonelinas (1994) conceptualized source memory as a pure recollection process, in that a correct response depends only on successful recollection of contextual details of the study event.

Given that source memory is dependent on recollection in this way, and that recollection according to the dual-process model is thresholded, the model implies that performance in source memory tasks should exhibit a retrieval threshold. When forced to make a source memory judgement in the case that recollection fails, participants can only guess in the absence of information. The dual-process model therefore predicts that overall performance on source memory tasks should be a mixture of two discrete components: informed responses and uninformed guesses, depending on whether recollection succeeds or fails respectively.

# Recognition and Source Memory in Two-choice Tasks

## Source ROC Predictions

Traditionally, evidence both for and against a threshold in recollection has come from the examination of Receiver Operating Characteristic (ROC) curves (Yonelinas & Parks, 2007; Yonelinas, 1999; Slotnick & Dodson, 2005). In a two-choice paradigm with two possible sources of information, each source in a continuous model is associated with a normally distributed memory strength, and these distributions overlap. As the response criterion is varied, the ratio of hit rates to false alarms will be such that the resultant shape of the plot is curvilinear (Slotnick & Dodson, 2005). In contrast, in a discrete model, the strength of the memory representation fails to meet either response threshold in the overlap. As a result, the ratio of false alarms to hit rates across criterion points is constant, producing a linear ROC (Rouder, Morey, Cowan, Zwilling, Morey & Pratte, 2008). The dual process model, in which source memory is dependent on a thresholded recollection process, also predicts linear source ROCs when familiarity is the same for both sources (Yonelinas, 1999).

### Conditioning Source Memory on Recognition Confidence

The premise that recollection is thresholded was challenged by a reanalysis of the Yonelinas (1999) data by Slotnick and Dodson (2005), in which they conditioned source performance on recognition confidence ratings for each item. This reanalysis demonstrated that if source ROCs were plotted separately for different levels of confidence reported in the item recognition task, the highest confidence source ROCs were in fact curvilinear, contrary to the predictions of the dual-process model. Performance for unrecognized items was at chance and these items were on the diagonal of the ROC. As items rated with lower recognition confidence were included in the original data, the shape of the overall source ROC were apparently increasingly linear, and more consistent with the predictions of the threshold model. The authors argued that only the items that were recognised with high confidence contained diagnostic source information, and that the linearity of source ROCs observed by Yonelinas (1999) was an artifact of collapsing across all recognition confidence ratings, and was thus not evidence for a recollection threshold.

Yonelinas and Parks (2007) responded to the Slotnick and Dodson (2005) analysis by proposing that source ROCs are typically linear, but become more curvilinear under a number of conditions. One such condition is when an item and a source are treated holistically as one item, known as *unitised familiarity*. We will return to this point in the Discussion. While this proposal represented a concession towards a continuous contribution under certain circumstances, Klauer and Kellen (2010) were later able to account for curvilinear ROCs using only discrete states by allowing for a variable mapping between recognition confidence ratings and source memory thresholds. At present, then, there is a lack of consensus about whether apparently linear or curvilinear ROCs reflect thresholded or continuous retrieval processes.

# Continuous Report

Harlow and Donaldson (2013) addressed the need for more diagnostic data by replacing binary decision outcomes with a continuous report paradigm, which yielded a continuous measure of response accuracy. In the Harlow and Donaldson (2013) continuous report paradigm, source information was provided by a point located on the circumference of a circle, which represented the context, and which was paired with a word item. When later cued with that word, participants were required to reproduce the associated location. This procedure allowed for a continuous measure of the error in the angle between the reported and true source locations. The researchers’ use of a continuous measure of source memory performance allowed them not only to measure the accuracy of source memory judgments, but also the distribution of response errors. Instead of categorizing responses as either correct or incorrect as in a two-choice task, their task, which captures an entire distributions of response accuracy, provides a more detailed picture of trial-to-trial variability in retrieval performance. The additional information in such distributions may be more diagnostic than ROC curves of the underlying retrieval processes. Critically, the threshold and continuous models of source memory make divergent predictions about the distributions of response errors in continuous report tasks.

According to the threshold model, items that fall below the recollection threshold will result in guesses, which will be distributed uniformly across all possible response options. Items that exceed the threshold and are successfully retrieved will cluster, with some error, around the true value of the item source. This two-process account of continuous report performance parallels similar proposals in the visual working memory literature, where Zhang and Luck (2008) used a two-component mixture model comprised of a von Mises distribution and a uniform distribution to argue for an item-capacity-limited model visual working memory. Items in memory are represented with high accuracy and responses to them follow a von Mises distribution; items not in memory lead to guessing and responses to them follow a uniform distribution. Harlow and Donaldson (2013) took a similar approach in modelling performance in their source memory task, using a von Mises distribution to characterize the shape of the marginal distribution of response errors when items exceeded the retrieval threshold. The von Mises distribution is a circular analogue of the normal distribution and, like the normal distribution, has a bell-shaped density function. A mixture of a von Mises distribution and a uniform distribution produces a high-peaked, heavy-tailed distribution (Harlow & Donaldson, 2013). Harlow and Donaldson (2013) found that source accuracy data was better fit by the threshold model better than by its continuous counterpart, which predicts that responses made with moderate memory strength would result in a wider spread of responses around the true location without a uniformly distributed guessing component.

## Source Memory for Unrecognized Items

Although Harlow and Donaldson’s (2013) method represents an innovative way to charactertize the retrieval processes in source memory tasks, a potential confound in their experimental design was that they did not condition source memory performance on recognition confidence. The Slotnick and Dodson (2005) study discussed earlier showed how source-memory ROC shapes depend on recognition confidence in the two-choice paradigm, and it is possible that continuous source memory judgments are affected in a similar way. Hautus, MacMillan, and Rotello (2008) modeled performance in two-choice source memory tasks, and found that the best fits were obtained from a model that incorporated a guessing process for unrecognized items. Their findings mirrors those of Slotnick and Dodson (2005) who found that source performance was at chance for items recognised with low confidence. Hautus et al. preferred a model with a guessing process because continuous models that lacked a guessing process predicted that the performance decreases that occurred with decreasing recognition confidence were too gradual and could not capture the abrupt shift to chance performance when items were unrecognized.

A lack of source discriminability for unrecognized items has been replicated numerous times (Bell, Mieth, & Buchner, 2017; Malejka & Broder, 2016; Onyper, Zhang, & Howard, 2010), although these studies often employed a procedure where item and source ratings were obtained in the same test trials. When item recognition and source memory tests were in separate blocks, Osth, Fox, McKague, Heathcote, and Dennis (2018) observed reliable source memory for unrecognized items, but discriminability was still quite low (*d’* ~ .1).

If the lack of source memory for unrecognized items generalizes to continuous report tasks, then guesses would result in a heavy-tailed error distribution, which would not necessarily reflect a threshold in memory retrieval but might simply reflect a state in which source retrieval was not attempted. Guessing behavior can arise either as the product of a retrieval threshold within source memory or without a threshold in source when either the participant does not attempt to retrieve the source memory or the source memory is absent. In the context of the findings of Harlow and Donaldson (2013), this latter account predicts that if unrecognized items are excluded, the heavy tails in the error distribution will disappear, and that a continuous model will be preferred in account for source performance. An aim of our study was therefore to investigate a continuous-report measure of source-memory performance conditional on the accuracy of previous recognition judgments.

# Decision-Making

In completing a source or recognition memory task, not only do participants need to retrieve information from memory, they must also make a decision on how to respond based on the information retrieved. In this sense, the information retrieved from memory can be thought of as evidence for entering a decision process. Much of the existing body of source memory research, particularly in the continuous report paradigm, lacks an explicit account of properties of the decision process. Past research in the recognition memory literature has shown that when the properties of decision processes are taken into account, , the kinds of conclusions that can be made about recognition memory differ from those made when decision-making is not explicitly considered (Dube, Starns, Rotello & Ratcliff, 2012; Osth, Bora, Dennis, & Heathcote, 2017; Ratcliff & Starns, 2013). Diffusion models have emerged as increasingly influential accounts of decision processes which predict both response time (RT) and response accuracy data, and which naturally explain well-documented phenomena like the speed-accuracy trade-off (Ratcliff, Smith, Brown & McKoon, 2016). Diffusion models have also been used extensively in the past to model memory retrieval, and more recent research has proposed a general theory of memory and decision-making in which decisions about stimuli within visual working memory are made using a diffusion process (Smith & Ratcliff, 2009). In the most common form of the diffusion model (Ratcliff, 1978), the decision process is modeled as noisy evidence accumulation between a pair of absorbing boundaries that represent the decision criteria for the task. Evidence is accumulated until the process reaches one or other boundary: The first boundary reached determines the response and the time to first reach a boundary is the decision time component of RT. The diffusion decision model in shown in Figure 1.

A close up of a map

Description automatically generated

Figure 1. Diffusion decision model. Evidence is accumulated by a Wiener diffusion between a pair of absorbing boundaries that represent the decision criteria for responses Ra or Rb. The starting point is z and the boundaries are located at 0 and a. The first boundary reached determines the response and the time taken to reach it determines the decision time. The rate at which evidence accumulates is the drift rate, which is normally distributed across trials with standard deviationη.

The diffusion model assumes that multiple sources of variability affect the decision process, including moment-to-moment variability in the accumulation of evidence and trial-to-trial variability in the quality of evidence entering the decision process. The moment-to-moment variability reflects the noisiness of the evidence provided by the retrieval process, while the trial-to-trial variability reflects differences in the stimulus information on which the decision is based. The rate of evidence accumulation on any trial is known as the drift rate. Drift rates can vary across trials, with high drift rate trials resulting in high accuracy and fast RTs, while trials with lower drift rates result in slower and less accurate responses (Ratcliff, Smith & McKoon, 2015).

The relationship between accuracy and RT is intuitive: when the evidence for the correct response is strong, it will accumulate more rapidly, leading to faster responses, and be more likely to reach the correct boundary before the error boundary, leading to higher accuracy. In contrast, when the evidence is weak, it will accumulate more slowly, leading to slower responses, and be less likely to reach the correct boundary before the error boundary, leading to lower accuracy. Consequently, when drift rate varies between trials, leading to a mixture of strong and weak evidence, then the mean RT for correct responses will be shorter than mean RT for errors, because correct responses are more likely to occur on trials in which driftrates are higher. This phenomenon, known as a *slow error* pattern,has been reliably observed when decision making is difficult (Ratcliff et al., 2016). A counterintuitive property of trial-to-trial variability in the drift rate of the diffusion model is that, without this variability, correct and error RT distributions will be the same, it is only when drift rates vary between trials that we obtain this slow error pattern.

The circular diffusion model, of Smith (2016) extended the two-choice diffusion model of Ratcliff (1978), which represents decision-making as a one-dimensional evidence accumulation process (diffusion on a line), to account for continuous report tasks. In the circular model, the drift rate is defined as a vector in a two-dimensional (2D) space having both an identity component for the position of the stimulus on the circle, represented as the phase angle of the drift vector, and a magnitude component, representing the quality of the evidence for any particular response (Figure 2). When a response is made, the magnitude of the drift vector determines RT in the same way as does the scalar drift rate does in the standard Ratcliff model, while the point at which the evidence accumulation process exits the circle determines the response outcome.

A close up of a device

Description automatically generated Xθ

Figure 2. Circular diffusion model of continuous report. Evidence is accumulated by a two-dimensional Wiener diffusion on the interior of a disk, whose bounding circle, of radius a, represents the decision criterion. Evidence is accumulated starting at the origin until the process hits the bounding circle. The hitting point, Xθ, is the decision outcome and the hitting time, Ta,is the decision time. The drift rate is vector-valued and consists of two components, (μ1, μ2), which jointly specify its magnitude and direction. In polar coordinates the magnitude is represented by the drift norm ||μ|| and direction is represented by the phase angle θμ The noisy sample path represents evidence accumulation on a single experimental trial.From P. L. Smith (20160. “Diffusion theory of decision making in continuous report’ Psychological review, 123, 425-451. Figure 2. Copyright American Psychological Association.

The properties of the circular diffusion model closely parallel those of the two-choice diffusion model. When the only source of variability in the model is moment-to-moment variability in the evidence accumulation process the model predicts that decision times will be the same for all decision outcomes. When there is across-trial variability in drift rates, the model predicts that accurate responses will be faster than inaccurate responses. When there is across-trial variability in decision criterion, represented by variability in the diameter of the bounding circle, the model predicts that accurate responses will be slower than inaccurate responses. These properties are continuous counterparts of the slow-error and and fast-error properties predicted by the two-choice diffusion model with across-trial variability in drift rate and starting point, respectively.

Mathematically, the circular diffusion model predicts that, for a fixed drift rate and decision criterion, the distribution of decision outcomes will follow a von Mises distribution. The spread of decision outcomes predicted by the von Mises distribution depends on a precision parameter, κ. Loosely, precision is the inverse of variance: high precision represents low variance and vice versa. The von Mises precision predicted by the circular diffusion model is jointly a function of the drift norm, ||μ||, the decision criterion, *a*, and the noise in the evidence accumulation process, σ2. Specifically,

In words, this equation says that precision is equal to the quality of the information in the stimulus, represented by the drift norm, multiplied by the amount of evidence required for a response, represented by the decision criterion, divided by the noisiness of the evidence accumulation process.

An important property of the circular diffusion model for this study is that, while a fixed drift rate predicts a von Mises error distribution, the introduction of trial-to-trial variability in drift rate leads to peaked, high-tailed distributions, similar to those found by Harlow and Donaldson (2013) and in the visual working memory literature (Zhang & Luck, 2008). The peaked high-tailed distributions are the result of mixing trials that have high and low drift rate norms. High and low drift rates lead to error distributions with high and low precision, respectively. Mixtures of high and low precision von Mises distributions lead to peaked, heavy-tailed distributions like those found experimentally (van den Berg, Awh, & Ma, 2014). Harlow and Donaldson interpreted these kinds of distributions as evidence of an underlying memory retrieval threshold.

# The Current Study

There are two aims in the current study. The first is to investigate whether the heavy-tailed distributions found by Harlow and Donaldson (2013) may be the result of source guessing for unrecognized items. We did this by investigating source memory performance conditional on confidence in the recognition task. If heavy-tailed distributions of errors are due to source guessing on unrecognized items, then they should be eliminated on trials on which recognition confidence is high.

Our experimental task also included a manipulation of the imageability and concreteness of the stimulus words, as rated on the MRC Psycholinguistic Database. Harlow and Donaldson (2013) selected words for low ratings on both metrics to prevent participants from visualizing a concrete object in a source location. In our study, we drew stimuli from pools of low and high imageability and concreteness words which allowed us to quantify and compare the effect of these attributes.

The second aim is to determine if incorporating a model of the decision-making process that allows for across-trial drift rate variability can capture the distribution of response error and RTs. If the heavy-tailed pattern in the Harlow and Donaldson (2013) data can be explained by across-trial variability in the drift rate of the circular diffusion model, then a threshold mechanism in the memory process is not necessarily implied by the heavy-tailed shape of the error distribution. Providing an elaborated model of decision-making in this manner offers additional constraint by jointly modeling response error and RTs. The across-trial drift rate variability account predicts a slow error pattern in the joint distribution.

# Method

## Stimuli and apparatus

Stimuli were presented on a 20’’ Dell 2009W LDC Monitor, set with a screen refresh rate of 60 Hz. Software written in MATLAB controlled stimulus presentation and recorded responses. Stimuli consisted of words generated from the MRC Psycholinguistic Database, selected/ for low concreteness (minimum 100, maximum 456) and imageability (minimum 100, maximum 481) in the low stimulus set, and high concreteness (minimum 543, maximum 611) and high imageability (minimum 545, maximum 609) in the high stimulus set. Words were displayed in size 24 point “Courier New” white font positioned in the center of a uniform mean luminance field.

## Participants

Twenty participants were recruited online through the University of Melbourne SONA system. Each participant was expected to complete four 60-minute sessions, for which they were paid $12 at the completion of each session. One participant who did not complete all four sessions was excluded from analysis (*N* = 19). All participants were provided with plain language statements and consent forms, and gave informed consent prior to data collection.

## Procedure

Participants completed the experimental tasks over four sessions, Each of the four sessions consisted of 180 trials, which was broken up into 18 blocks of 10 items each. Blocks were comprised of a study phase, followed by a test phase (Figure X).

In the study phase, participants were presented with a black cross on a dark gray outline at the start of each trial for 600 ms, which was followed by the display of a word in the centre of the screen for 1500 ms. To ensure that participants attended to the source information, they were regarded to indicate the previous location of the cross on the blank target circle using a computer mouse. Responses made within 6 degrees of the true target location were classified as attended and advanced participants to the next item. Responses further away were deemed unattended and the words “TRY AGAIN” was displayed for 1000 ms, then the location was then re-presented for 250ms, and the verification task was repeated. Participants were then instructed to complete a distractor task, which involved 30 seconds of arithmetic problems. Following this, participants were shown a scrambled list of 10 previously studied items and 10 foils, and asked to rate each item on a six-point confidence Old/New scale. Finally, in the source memory retrieval task, participants were cued with the words for 1500 ms, and then indicated the recalled location by a clicking a mouse on the circumference of a grey response circle. There was no time limit on the decision task. A schematic for one trial in each of the phases is shown in Figure 3.

A picture containing screenshot

Description automatically generated

Figure 3. Schematic of display presented to the participant in one trial in each phase of the experiment. There was also an arithmetic distractor task between the encoding and recognition phases of the block, which is not depicted in the figure. The mouse cursor is shown in the center of the “Source Task” panel to illustrate the procedure. In the actual experiment, cursor was hidden from the participant and replaced with a red dot with a diameter of four pixels.

# Results

The results are presented in five parts. First, we ascertain whether individual participants’ responses in the source retrieval task was above chance. As responses were continuous, above-chance performance translates into a deviation from uniformity in responding. We did this in order to distinguish participants who were responding at chance from those who showed better-than-chance source memory performance. Second, we investigate source memory judgments conditioned on the prior recognition response and show that conditioning source responding on successful recognition does not fully account for the heavy tails in the distribution of source memory accuracy. Third, we fit the Zhang and Luck (2008) uniform-plus-von Mises mixture model to the marginal distributions of response accuracy, conditioned on recognition performance, and show that recognition affects the precision of the source information that is retrieved and not the proportion of guessing responses. Fourth, we present fits of versions the circular diffusion model to the joint distributions of RT and accuracy. Fifth, we present a more general version of the circular diffusion model in which the phase angles of the drift rates, which represent the identities of the items retrieved from memory, are represented by a nonnormal distribution that we call a *generalized von Mises distribution*. We show that the fits of the various versions of the model converge on a shared conclusion that source responding is subject to a retrieval threshold.

**Data Screening**

Preliminary inspection of the data suggested that some participants performed the source retrieval task with very low accuracy. A Rayleigh test for uniformity identified two participants whose data did not indicate evidence for a departure from uniformity in at least one condition, interpretable as completely random responding (Table X; Fisher, 1993). These participants will be referred to as a *low response accuracy* subgroup, with the expectation that the data from the remaining *high response accuracy* group will be more diagnostic for the purposes of distinguishing between the models.

|  |  |  |
| --- | --- | --- |
| Table X  Rayleigh Test for Uniformity for Source Memory Response Error | | |
| Participant | Mean Resultant Length | *p* |
| 1 | 0.02 | .87\* |
| 2 | 0.69 | <.01 |
| 3 | 0.44 | <.01 |
| 4 | 0.51 | <.01 |
| 5 | 0.55 | <.01 |
| 6 | 0.21 | <.01 |
| 7 | 0.57 | <.01 |
| 8 | 0.87 | <.01 |
| 9 | 0.10 | .01 |
| 10 | 0.87 | <.01 |
| 11 | 0.35 | <.01 |
| 12 | 0.66 | <.01 |
| 13 | 0.07 | .24\* |
| 15 | 0.08 | .04 |
| 16 | 0.54 | <.01 |
| 17 | 0.09 | <.01 |
| 18 | 0.62 | <.01 |
| 19 | 0.29 | <.01 |
| 20 | 0.28 | <.01 |

\* *p* values greater than 0.05, indicating no evidence of a departure from uniformity for participants 1 and 13.

## Source Memory for Unrecognized Items

The data for each participant were split into three categories on the basis of participants’ confidence in the recognition phase of the experiment. Items which were rated three and below were deemed unrecognized by the participants, while successful recognition was defined by a rating of four and above. Of the recognized items, ratings of the maximum value of six are further specified as highly recognized items.

The results of Rayleigh tests of uniformity of source responses for unrecognized items are displayed in Table X. The distributions of these responses were uniform for all participants, indicating that no source memory was present when recognition confidence was low.

|  |  |  |
| --- | --- | --- |
| Table X  Rayleigh Test on Source Memory for Unrecognized Items | | |
| Participant | Test Statistic | *p* |
| 1 | 0.06 | 0.46 |
| 2 | 0.24 | 0.38 |
| 3 | 0.17 | 0.16 |
| 4 | 0.11 | 0.48 |
| 5 | 0.04 | 0.96 |
| 6 | 0.13 | 0.30 |
| 7 | 0.08 | 0.75 |
| 8 | 0.33 | 0.17 |
| 9 | 0.11 | 0.80 |
| 10 | 0.38 | 0.07 |
| 11 | 0.40 | 0.01 |
| 12 | 0.18 | 0.40 |
| 13 | 0.10 | 0.62 |
| 15 | 0.33 | 0.02 |
| 16 | 0.20 | 0.06 |
| 17 | 0.05 | 0.76 |
| 18 | 0.37 | 0.02 |
| 19 | 0.21 | 0.17 |
| 20 | 0.19 | 0.44 |

**Item Recognition Performance**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Condition | Hit Rate | | False Alarms | |
|  | *M* | *SD* | *M* | *SD* |
| High Imageability | .87 | .14 | .14 | .10 |
| Low Imageability | .84 | .13 | .14 | .11 |

An independent samples *t*-test applied to individual-level hit rates for high and low imageability conditions indicated that there was no significant difference in hit rates across the two conditions *t*(35.85) = .68, *p* = .503. Coupled with the Rayleigh test on unrecognized items, this suggests that although source performance for unrecognized items was uniform, the majority of items were successfully recognized, and so guessing due to recognition failure does not fully account for the heavy tails of the error distributions. This can be confirmed visually by comparing histograms constructed from group-level source responses conditioned on recognition rating in Figure X.

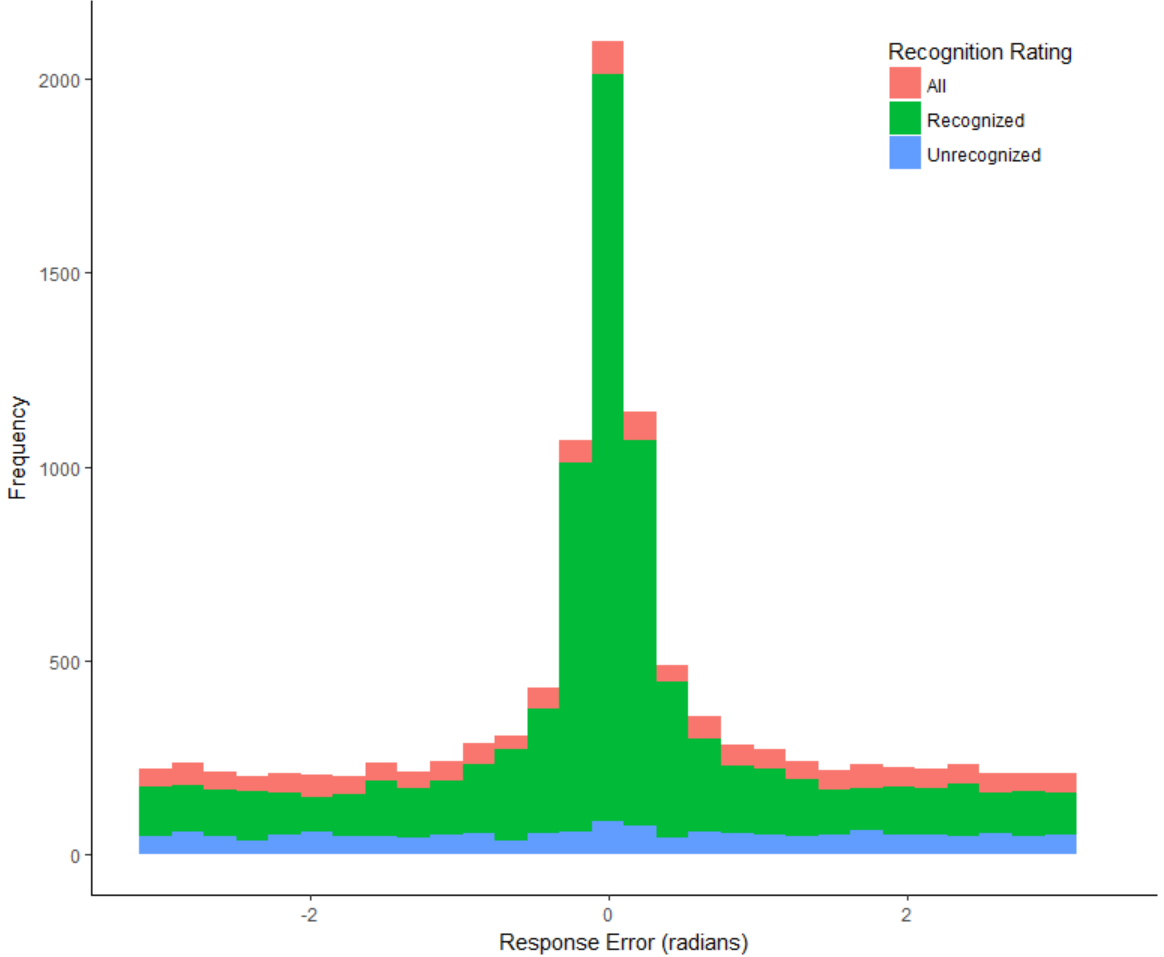


Figure 4. Frequency of angular response error in the source retrieval task, collapsed across participants. The subset of unrecognized items (rated three or below on the six-point confidence scale) yields source responses which are uniform, indicating no source memory for unrecognized items. However, the frequency of unrecognized items was relatively low and the exclusion of these items from the distribution of responses (in red), does not eliminate the heavy tails from the distribution of responses for recognized items (in green).

**Simple Mixture Model**

To attempt to replicate the Harlow and Donaldson (2013) finding, we first used the Zhang and Luck (2008) mixture model to fit the marginal distribution of response error. The model had two free parameters, one for the von Mises precision, which described the spread of responses around the true location, and a mixing parameter *π*,which described the proportion of trials which were driven by information in a von Mises distribution, as opposed to guesses in a uniform distribution. To clearly distinguish this model from mixture variants of the circular diffusion model, this will be referred to as the Simple Mixture model.

The mean best fitting parameters of the Simple Mixture model to the response accuracy data, excluding the low response accuracy group, are shown in Table X. The parameter estimates at an individual level are in Appendix X. Conditioning on recognition largely affects the precision parameter and not the memory parameter.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Table X | | | | |
| *Parameter Values for Best Fits of the Simple Mixture Model to Highly Recognized Individual Data.* | | | | |
| Recognition Rating | Low Imageability | | High Imageability | |
|  | Precision | *π* | Precision | *π* |
| All | 19.22 | 0.51 | 23.89 | 0.51 |
| Recognized | 18.24 | 0.50 | 23.79 | 0.54 |
| Highly Recognized | 18.03 | 0.50 | 19.81 | 0.56 |

*Note.*  Precision refers to the precision of the information-driven retrieval process. *π* represents proportion of responses driven by information.

**Circular Diffusion Models**

We tested three alternative versions of the circular diffusion model that expressed different hypotheses about the process of memory retrieval. The first of these was designed to be analogous to the continuous model of source memory presented in Harlow and Donaldson (2013), and was implemented as a circular diffusion model with across-trial variability in drift rates. This variant will subsequently be referred to as the *continuous diffusion* model. Drift rate variability was set to be equal in both dimensions of the two-dimensional (2D) space, but different between imageability conditions. There were four parameters for mean drift rates (*μ1x*, *μ2x*, *μ1y*, *μ2y*),which characterize the components of the drift rate in the *x* and y directions in the low (*μ1*) and high (*μ2*) imageability conditions. Because the model was fitted to the distribution of report errors, which is centered on zero degrees, the dominant component of the drift rate was expected to be in the *x* direction, which corresponds to a phase angle of zero). The second component of drift was included to allow for the possibility of drift bias. The decision criterion was represented by *a*, which had uniform variability across trials with range *sa*. There were also two standard deviation parameters *η1* and *η2*, which described the standard deviations of the drift rates in the low and high imageability conditions respectively. The standard deviations of the drift rates were assumed to be the same in the *x* and *y* directions. Finally, there was a non-decision time parameter, *Ter­*. Like the standard diffusion model, the circular model assumes that RT is the sum of the decision time and a time for other (encoding and response) processes. These parameters are summarized in Table 2.

The second model variant embodied the thresholded retrieval property preferred by Harlow and Donaldson (2013), and will be referred to as the *threshold diffusion* model. This was implemented as a mixture of two diffusion processes: one with positive drift rate and no between-trial drift variability, and a second that was modeled as a diffusion process with zero drift rate. The zero-drift process provides a diffusion process implementation of a guessing process, in which the decision process is driven only by noise. Unlike “guessing” in its classical sense, which accounts for accuracy but not RT, the zero-driven process is able to predict both accuracy and RT. Mixing proportions for the two processes were allowed to vary between the imageability conditions. This model had ten free parameters. Four mean drift rates parameters were shared with the continuous model (*μ1x*, *μ2x*, *μ1y*, *μ2y*), with the same interpretation, as well as *Ter­,* the non-decision time parameter. There were two parameters for the mixing proportions between information-driven and guessing components, one for the low imageability condition (*π1*) and another for the high imageability condition (*π2*). The decision criterion was estimated separately for the information-driven component (*a1*) and the guessing component (*a2*). Both processes shared a parameter for criterion variability (*sa*).

The third model was a combination of the continuous and threshold diffusion models. It assumed a mixture of zero-drift and nonzero-drift processes, like the threshold diffusion model, but also allowed for across-trial variability in drift rates This model can therefore be regarded as a mixture model that incorporates both the continuous and threshold diffusion models. This, the *mixture diffusion* model, had 11 free parameters, all of which are displayed in Table X.

|  |  |
| --- | --- |
| Table X  Symbols and definitions of free parameters estimated in diffusion model variants | |
| Symbol | Parameter |
| *μ1a* | Mean drift, low condition, x direction |
| *μ2a* | Mean drift, low condition, y direction |
| *μ1b* | Mean drift, high condition, x direction |
| *μ2b* | Mean drift, high condition, y direction |
| *η1* | Drift variability, low condition |
| *η2* | Drift variability, high condition |
| *a1* | Decision criteria, information-driven component |
| *a2* | Decision criteria, guessing component |
| *π1* | Mixing proportion, low condition |
| *π2* | Mixing proportion, high condition |
| *Ter* | Non-decision time |
| *sa* | Criterion variability |
| *st* | Non-decision time variability |

*Note.* Not all parameters were estimated for all three models. The continuous diffusion model did not include a mixed guessing process, and therefore lacked *a2*, *π1* and *π2.* The threshold diffusion model did not have drift variability and lacked *η1* and *η2*. The mixture diffusion model included all parameters listed.

The three variants of the circular diffusion model were each fit to data at an individual level. The Bayesian Information Criterion (BIC) and Log Likelihood (LL) for the three models’ fits to each participant is shown in Table X.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Table X | |  | |  | |  | |
| Bayesian Information Criterion (BIC) and Log Likelihoods (LL) for Fits of the Models to Individual Data | | | | | | | |
| Participant | | Continuous | | Threshold | | Mixture | |
| BIC | LL | BIC | LL | BIC | LL |
| High Precision | 2 | 1072.10 | 2200.60 | 1003.11 | 2068.90 | **942.03** | 1959.27 |
|  | 3 | 1003.21 | 2061.94 | **851.91** | 1765.53 | 855.20 | 1784.44 |
|  | 4 | 1907.85 | 3878.09 | 1702.98 | 3475.27 | **1616.18** | 3315.53 |
|  | 5 | 1127.15 | 2310.82 | **1014.05** | 2090.91 | 1127.19 | 2329.75 |
|  | 6 | 771.93 | 1599.33 | **747.05** | 1555.73 | 751.15 | 1576.26 |
|  | 7 | 766.43 | 1588.75 | **648.91** | 1359.93 | 1529.70 | 3133.92 |
|  | 8 | 560.81 | 1178.03 | **302.98** | 668.64 | 584.20 | 1243.61 |
|  | 9 | 1477.13 | 3010.53 | 1467.42 | 2997.37 | **1454.16** | 2983.34 |
|  | 10 | 1023.30 | 2103.01 | 1044.56 | 2151.79 | **994.37** | 2063.94 |
|  | 11 | 1244.35 | 2544.86 | 1103.18 | 2268.76 | **901.70** | 1878.29 |
|  | 12 | 743.68 | 1543.59 | 493.95 | 1050.38 | **476.26** | 1027.49 |
|  | 15 | 1300.51 | 2657.06 | 1343.38 | 2749.03 | **936.71** | 1948.13 |
|  | 16 | 1000.38 | 2056.27 | 894.16 | 1850.00 | **888.14** | 1850.28 |
|  | 17 | 1145.93 | 2045.93 | **1012.46** | 2085.01 | 1039.42 | 2150.94 |
|  | 18 | 1030.68 | 2117.71 | 2049.63 | 4161.88 | **884.49** | 1844.12 |
|  | 19 | 1782.81 | 3621.04 | 1469.25 | 3000.09 | **1391.70** | 2857.30 |
|  | 20 | 1590.44\* | 3236.77 | 1666.55 | 3395.18 | 1685.63 | 3445.76 |
|  |  |  |  |  |  |  |  |
| Low Precision | 1 | 574.31 | 1197.40 | 635.49 | 1325.19 | **573.01** | 1211.06 |
|  | 13 | 2079.41 | 4214.00 | 1717.02 | 3495.35 | **1558.26** | 3190.09 |

\* Indicates lowest BIC

Both the threshold and the mixture models consistently outperformed the continuous model without guessing. This strongly suggests that participants sometimes do respond in a no-information guessing state, which is mixed with a distribution of responses driven by information which is centered on the target location. In comparing the two models which utilize a threshold, the mixture model appears to fit the data of most participants better than the pure threshold model, but this advantage is very slight and is outweighed by the penalty for complexity applied by the BIC for the two additional parameters allowing for trial-to-trial variability, as shown by the mixture model having the lowest negative log likelihood (LL) and the threshold model having the lowest BIC for most participants (Table X). This suggests that the addition of drift variability does not improve the fit of the threshold model enough to justify the additional complexity introduced into the model.

**Generalized von Mises Model**

Schurgin, Wixted, and Brady (2018) recently proposed that the heavy-tailed distributions of errors found in visual working memory studies could be accounted for by an exponential-like compression of the psychological space, which makes pairs of stimuli that are far from a given reference stimulus appear closer together than they actually are. They found that a combination of a signal detection decision model and a nonlinear transformation of the psychological space could predict heavy-tailed distributions like those commonly found in visual working memory and predicted by Zhang and Luck’s (2008) mixture model. Schurgin et al.’s model is a model of response accuracy only, and makes no predictions about RT. However, we can incorporate similar ideas into the circular diffusion model if we assume that the across-trial distributions of drift rates are nonnormal in form. Specifically, we assumed that the distribution of drift rate phase angles, which represents the retrieved stimulus identities across trials, had a flexible form, which we called a generalized von Mises distribution. This distribution allows for the possibility that, on a proportion of trials, the retrieved source information may be very inaccurate.

We use a version of the circular diffusion model, which draws upon a generalized von Mises distribution of phase angles, to fit data in the same source memory task. This model will be referred to as the *generalized von Mises* model.

|  |  |
| --- | --- |
| Table X  Symbols and definitions of parameters in the Generalized von Mises Model | |
| Symbol | Parameter |
| *μ1* | Mean drift rate vector length, low imageability |
| *μ2* | Mean drift rate vector length, high imageability |
| *κμ1* | Precision, low imageability |
| *κμ2* | Precision, high imageability |
| *η* | Drift variability |
| *ρ* | Distribution of phase angle |
| *a* | Decision criterion |
| *sa* | Decision criterion variability |
| *Ter* | Non-decision time |
| *st* | Non-decision time variability |

The critical parameters in the model are *ρ* and *κμ*,which determine the shape of the generalized von Mises distribution of drift rate phase angles. When *ρ* = 1, the distribution described in Smith et al. (2019) is equivalent to a standard von Mises distribution. As *ρ* approaches zero, the shape of the distribution becomes increasingly leptokurtic, such that for very low values of *ρ* (e.g., *ρ* = .05), the distribution resembles a spike of probability mass superimposed on a uniform background.

The interaction between various values of *ρ* and *κμ* produces model variants which can be related to the competing theories of source memory. When *ρ*=1 and *κμ* is large, the model resembles a fixed precision model in VWM. As *ρ* decreases, the model becomes increasingly similar to either a variable-precision model, based on the value of *κμ*.

A picture containing text

Description automatically generated

Ultimately, this produces a flexible model that is able to gradually transition to resemble models like the continuous and threshold circular diffusion models presented earlier, simply by adjusting parameters that govern the shape of the distribution of phase angles.

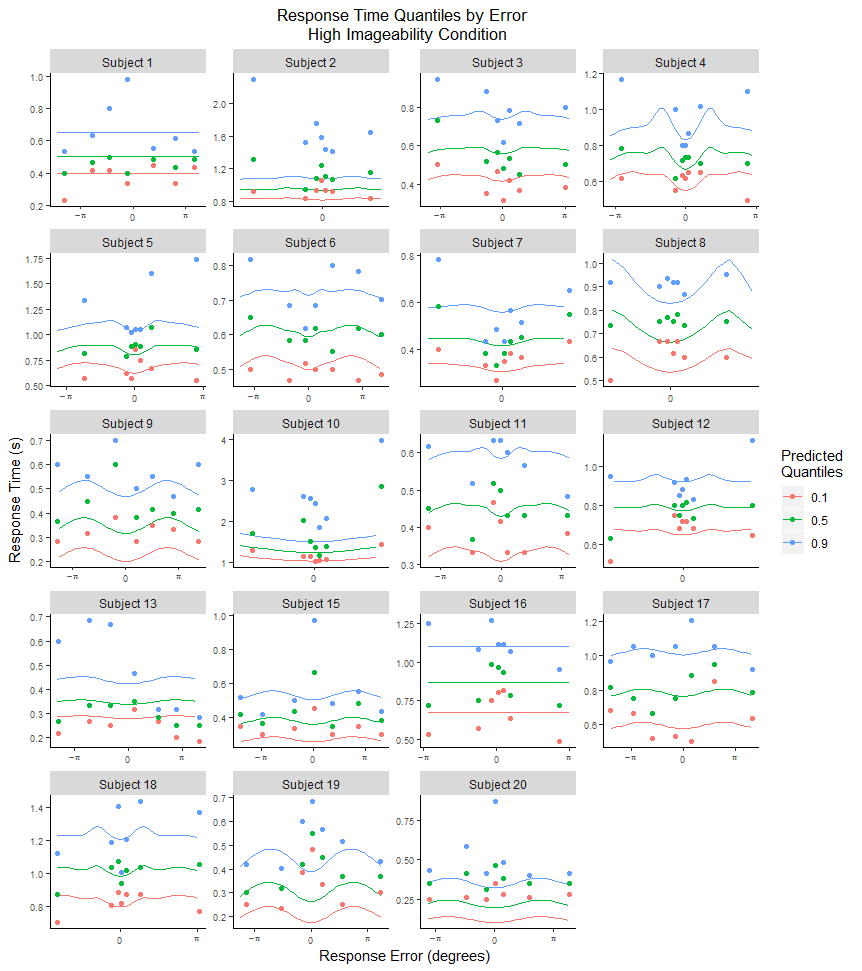
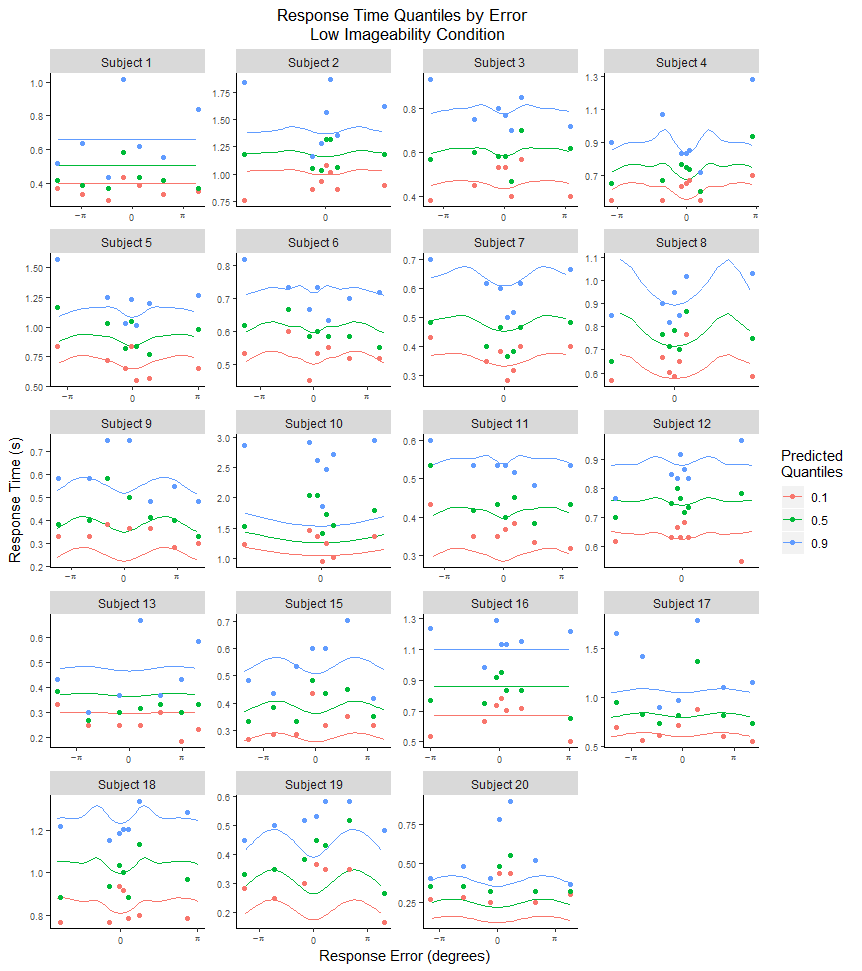


Figure X. Fit of the Generalized Von Mises model



All parameter estimates for each participant that generated the best fits to data are displayed in Table X.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table X  Parameter Estimates for Best Model Fits of the Generalized von Mises model without Criterion Variability | | | | | | | | | |
| Participant | *μ1* | *μ2* | *κ1* | *κ2* | *η* | *ρ* | *a* | *Ter* | *st* |
| 1 | 3.40 | 3.38 | 0.39 | 0.61 | 1.47 | 0.20 | 2.67 | -0.34 | 0.11 |
| 2 | 2.96 | 3.04 | 3.99 | 4.31 | 0.51 | 0.06 | 4.51 | -0.32 | 0.00 |
| 3 | 3.94 | 3.96 | 2.77 | 2.95 | 1.17 | 0.16 | 3.73 | -0.35 | 0.11 |
| 4 | 4.42 | 4.50 | 3.16 | 3.57 | 1.35 | 0.02 | 4.27 | -0.17 | 0.02 |
| 5 | 3.70 | 3.85 | 3.46 | 3.81 | 1.18 | 0.09 | 4.50 | -0.30 | 0.09 |
| 6 | 1.11 | 1.66 | 10.53 | 9.26 | 1.28 | 0.02 | 1.06 | 0.22 | 0.00 |
| 7 | 3.67 | 3.67 | 3.73 | 3.81 | 0.80 | 0.07 | 2.63 | -0.33 | 0.30 |
| 8 | 3.89 | 3.93 | 4.73 | 5.69 | 0.90 | 0.03 | 3.54 | -0.10 | 0.00 |
| 9 | 1.85 | 2.03 | 0.48 | 0.10 | 0.37 | 0.31 | 1.45 | -0.16 | 0.05 |
| 10 | 3.35 | 3.44 | 5.37 | 6.40 | 1.30 | 0.02 | 4.52 | 0.00 | 0.14 |
| 11 | 0.91 | 0.96 | 6.79 | 10.50 | 1.18 | 0.04 | 1.22 | -0.06 | 0.01 |
| 12 | 3.97 | 3.81 | 3.98 | 3.85 | 0.01 | 0.08 | 4.51 | -0.35 | 0.06 |
| 13 | 1.63 | 2.13 | 1.31 | 0.40 | 1.46 | 0.02 | 1.09 | -0.16 | 0.15 |
| 15 | 3.34 | 3.76 | 1.18 | 0.57 | 1.94 | 0.13 | 1.56 | -0.14 | 0.33 |
| 16 | 2.44 | 2.45 | 3.32 | 3.87 | 0.39 | 0.10 | 3.07 | -0.22 | 0.02 |
| 17 | 2.49 | 2.60 | 0.61 | 0.87 | 0.56 | 0.07 | 3.25 | -0.35 | 0.08 |
| 18 | 3.55 | 3.60 | 3.71 | 3.55 | 0.96 | 0.10 | 4.51 | -0.22 | 0.08 |
| 19 | 3.41 | 3.41 | 1.76 | 2.12 | 0.82 | 0.09 | 2.16 | -0.31 | 0.19 |
| 20 | 0.38 | 1.35 | 5.47 | 2.22 | 0.62 | 0.17 | 1.05 | -0.14 | 0.00 |

The values for the critical parameters *ρ* and *κ* for the high imageability condition (with *κ* designated by *κ2*) are plotted against each other in Figure X. Barring two participants, *ρ* was generally close to zero across participants, indicating that a leptokurtic distribution of drift rates was preferred when fitting the model to response error data. The precision of the diffusion process, *κ*, varied to a greater degree between participants.

A picture containing screenshot

Description automatically generated

Figure X.Values for parameters *ρ,* which governs the shape of the distribution of drift rates, and *κ*which governs precision.

# Discussion

In this article, we had two main aims, one theoretical and one empirical. Our theoretical aim was to attempt to characterize performance on a continuous-report source memory task using a mathematical model of the decision process, the circular diffusion model, to ascertain whether it could predict the distributions of decision outcomes and RT from such a task. In applying the model to this kind of task, we sought to ascertain whether the conclusions of Harlow and Donaldson (2013), which were based on the distributions of decision outcomes only, would continue to hold when both outcomes and RTs were taken into account. Our empirical aim was to ascertain whether Harlow and Donaldson’s (2013) conclusion that source memory is thresholded would continue to hold for memory when conditioned on item recognised with high confidence.

We found three pieces of convergent evidence that source memory retrieval is indeed best characterized as a thresholded process. Firstly, we found that even when source responses were conditioned on successful recognition, the marginal distribution of response error was well characterised with the Zhang and Luck (2008) mixture model, consisting of a von Mises and a uniform distribution. This corroborates the Harlow and Donaldson (2013) finding which used a wrapped Cauchy to similarly account for a greater number of very high accuracy and very low accuracy responses, with fewer responses with moderate accuracy than in a wrapped normal distribution. Secondly, in fulfilling our theoretical aim of predicting joint distributions of source response error and RT using the circular diffusion model, we found that the threshold and hybrid models, which both assumed a mixture of guessing and memory-based responses, fit the data better than the continuous model which did not. Thirdly, with the generalized von Mises model, we allow for non-linear scaling of source location stimuli to the representation of this information which serves as the phase angle component of drift in the evidence accumulation process. The earlier comparison of circular diffusion models suggests that a threshold underlies performance in source memory tasks. The generalized von Mises model arrives at a similar conclusion through an entirely different parameterization of across-trial variability.

**Response Time Modelling**

In this paper we present a novel application of the circular diffusion model to memory data. (Van den berg, 2014) plateaus in memory, such as the uniform distribution in a response error distribution, are statistically meaningless. Looking at RT distribution alleviates this by providing richer data for more constrained models. Because the model has to account for both jointly, including RT distributions in the analysis becomes a more stringent test of the model.

**Implications for Models of Memory**

In our study, jointly modelling RT along with response error provided evidence for a thresholded model of source memory. Our findings fall within the growing body of work in memory research that suggests the architecture of memory involves a memory strength threshold. Retrieval from memory fails discretely when strength is below this threshold, but precision is variable for responses where retrieval was successful (Onyper, Zhang, & Howard, 2010). A long-standing debate in the working memory literature is whether memory capacity is better described as a number of discrete slots or by a continuous shared resource (Bays, Catalao, & Husain, 2009; Luck & Vogel, 1997; Pashler, 1988; Wilken & Ma, 2004; Zhang & Luck, 2008). Adam, Vogel and Awh (2017) found evidence for a discrete item limit, using a whole report paradigm in which participants performed increasingly worse over three to four items in a six item list, and subsequent responses were best characterized as guesses. In the change detection literature, Donkin and Nosofsky (2016) found RT evidence supporting a hybrid model of visual working memory consisting of a mixture of guessing and memory-driven states. These models fall in a general family of models akin to the Zhang and Luck (2008) “slots + resources” model where a continuous resource is divided among slots, much like performance in our task can be described as variable but thresholded.

**Applications of the Circular Diffusion Model**

This study presents the first application of the circular diffusion model to continuous report memory data. The model was successful in providing a quantitative account of both RTs and response accuracy. Continuous report tasks are more informative than two-choice alternatives because they allow insight into the precision with which a response is made, rather than categorizing responses as correct or incorrect. The circular diffusion model is able to capitalize on the additional information continuous report affords to investigate the properties of decision-making in cognitive tasks. By characterizing the decision processes when response outcomes are continuous, the circular diffusion model represents the latest development in tools with which to investigate cognitive processes underlying performance, through understanding the properties of decision-making that ultimately translate cognition into a response.

**Fast Error Pattern in Group Level Data**

When we analysed group level data, we found a fast error pattern in the joint distribution of RTs and error (i.e. less accurate responses were associated with shorter RTs). The diffusion model is able to account for a fast error pattern by allowing trial-to-trial variability in the criterion (a) of the diffusion process, and so we incorporated this into all of the models. At an individual level however, there was no fast error pattern (refer to flat quantiles in quantile-quantile plot).

**Imageability**

**Individual Differences**

Participants varied greatly in how they responded to the task. There appear to be three distinct groups of participants based on responses to the task. Some participants (1, 9, 13, 15) have very low memory precision (response error for 1 and 13 don’t significantly deviate from uniformity), and are not very diagnostic of model performance. On the other extreme, participant 10 is also qualitatively distinct from the rest of the participants, with a pattern of extremely precise responses with very long RTs (on the order of 3-5s- compare medians?). For this participant, all models struggle to account for the peak of the response error distribution, because they are trying to simultaneously capture the slow RTs. Performance also varies between the remaining 14 participants, but they are at least engaging with the task at similar latencies and with non-zero precision.

**Future Work**

**Sequential Presentation of Stimuli**

Some qualifications must be made when drawing conclusions from this study. Firstly, the sequential presentation of item and source information may constitute a methodological bias towards the appearance of discrete failures as the temporal separation of the two parts increases the difficulty of binding the item to its supposedly associated source. The current modelling exercise is not able to distinguish between errors arising from such a failure and errors due to a retrieval threshold. The original motivation for presenting item and source information in this manner was to replicate the Harlow and Donaldson (2013) paradigm, in which these components were separated to prevent *unitized familiarity*, meaning that unitization of source and item might allow participants to use familiarity to complete the task, making it “more difficult to isolate a recollection threshold”. This methodology assumes a Yonelinas (1999) dual-process framework, and potentially biases results to reflect the supposed recollection threshold it was meant to isolate. There is no reason to expect that source memory in a natural environment would operate under these conditions, and a model of source memory should be able to characterize performance when source and item information is presented simultaneously. A useful future experiment would be to modify the experimental paradigm so that source and item are presented simultaneously to investigate if the models presented in this article perform similarly under these circumstances.

**Generalizing to other source modalities**

Work in the source memory literature has used a variety of source modalities aside from location. This includes color, the gender of the voice of a speaker.

**Conclusion**

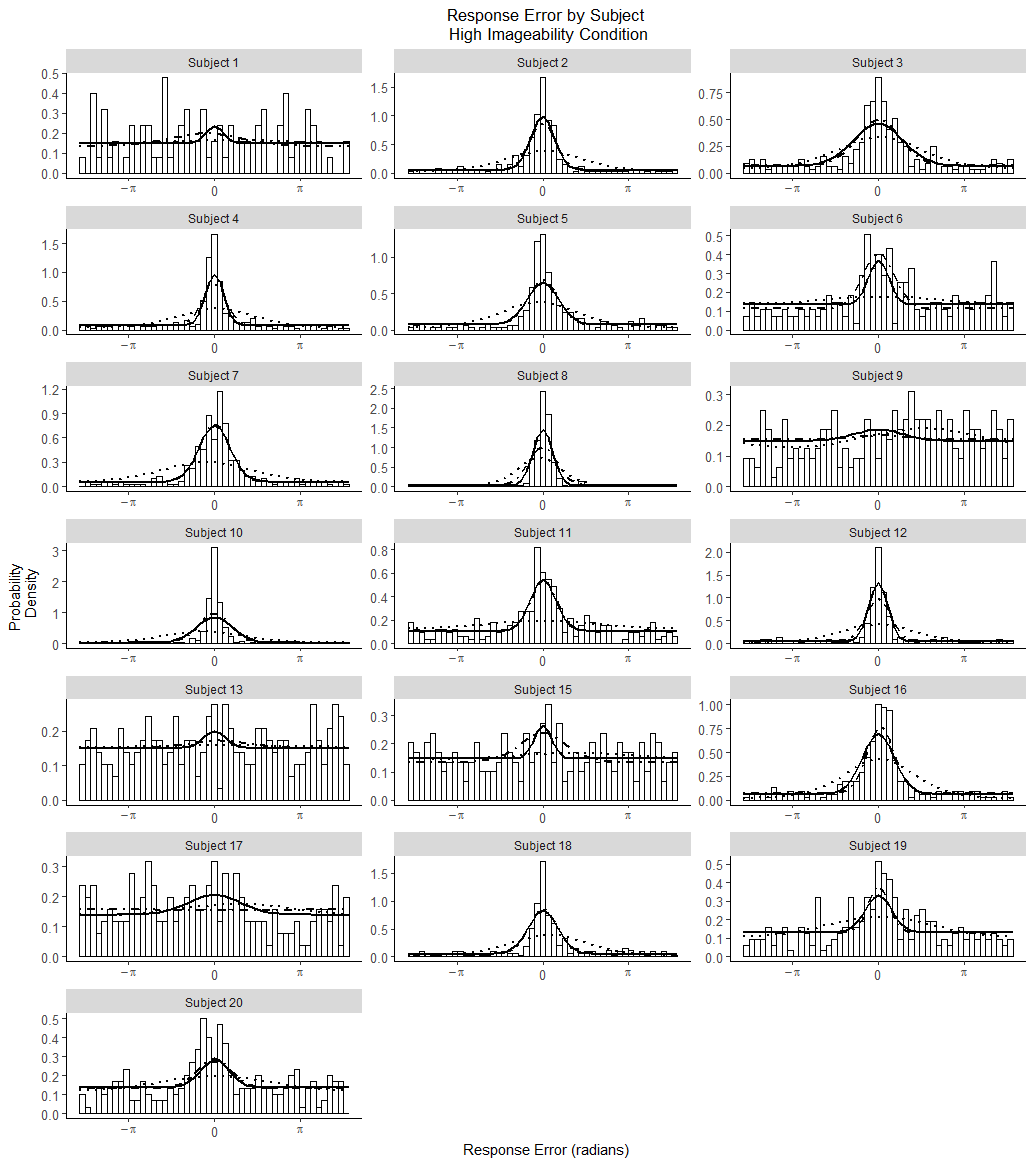
**Appendices**

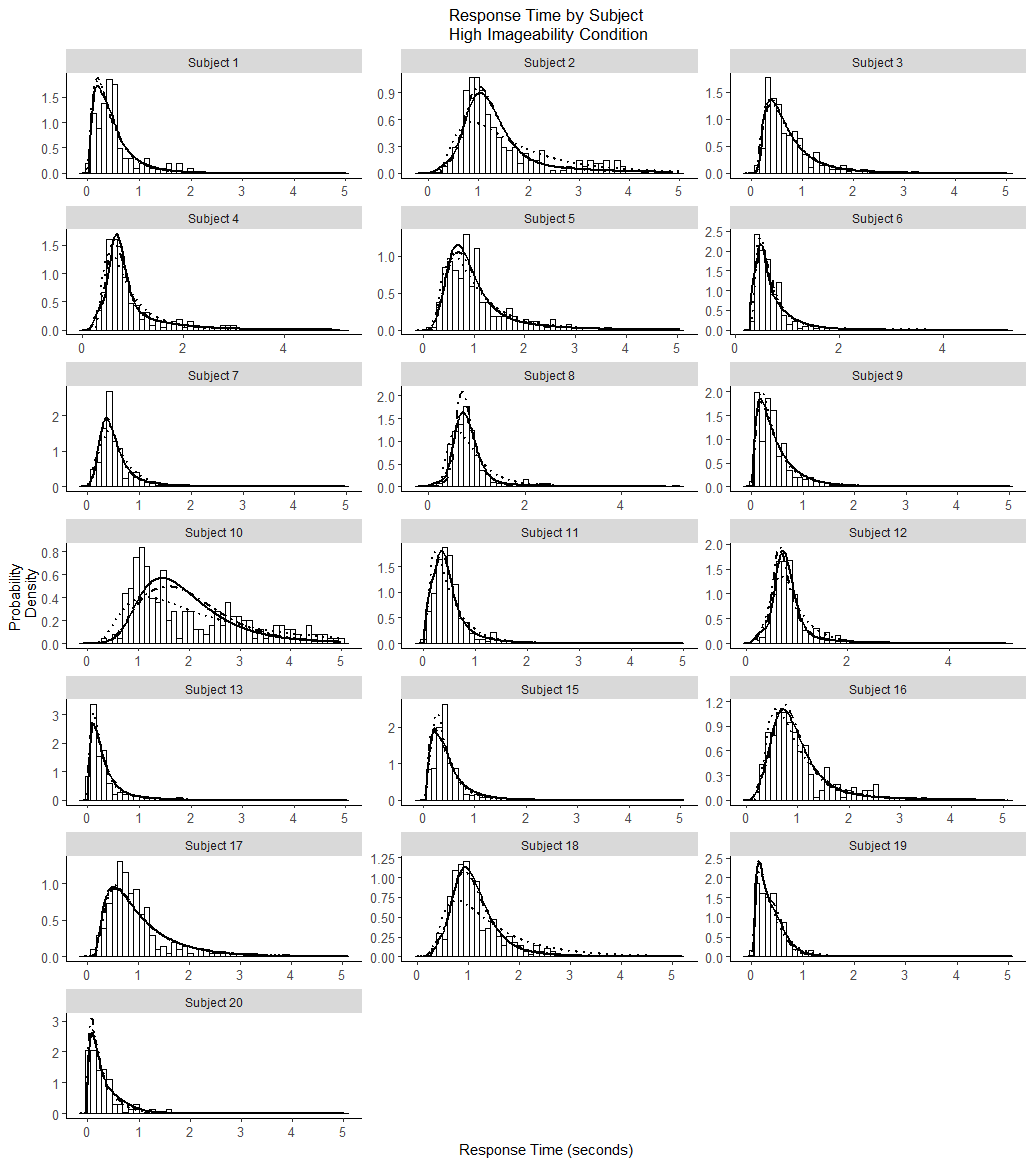
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Table X | | | | | |
| *Parameter Values for Best Fits of the Simple Mixture Model to All Individual Data.* | | | | | |
|  | | Low Imageability | | High Imageability | |
| Participant | | Precision | *π* | Precision | *π* |
| High Accuracy |  |  |  |  |  |
|  | 2 | 25.09 | 0.63 | 17.12 | 0.71 |
|  | 3 | 10.68 | 0.39 | 10.66 | 0.51 |
|  | 4 | 31.42 | 0.45 | 46.22 | 0.50 |
|  | 5 | 20.65 | 0.45 | 18.52 | 0.61 |
|  | 6 | 15.63 | 0.20 | 13.41 | 0.17 |
|  | 7 | 12.88 | 0.59 | 8.93 | 0.63 |
|  | 8 | 40.57 | 0.80 | 43.17 | 0.89 |
|  | 9 | 0.22 | 0.77 | 0.07 | 1.00 |
|  | 10 | 53.96 | 0.81 | 51.78 | 0.85 |
|  | 11 | 8.69 | 0.32 | 13.83 | 0.38 |
|  | 12 | 37.94 | 0.64 | 44.94 | 0.64 |
|  | 15 | 0.20 | 1.00 | 82.47 | 0.03 |
|  | 16 | 16.50 | 0.51 | 11.31 | 0.58 |
|  | 17 | 0.99 | 0.11 | 4.62 | 0.10 |
|  | 18 | 10.03 | 0.66 | 25.56 | 0.58 |
|  | 19 | 13.73 | 0.17 | 3.34 | 0.33 |
|  | 20 | 27.57 | 0.13 | 10.19 | 0.22 |
|  |  |  |  |  |  |
| Low Accuracy | 1 | 10.89 | 0.04 | 250.00 | 0.02 |
|  | 13 | 64.12 | 0.02 | 249.95 | 0.01 |

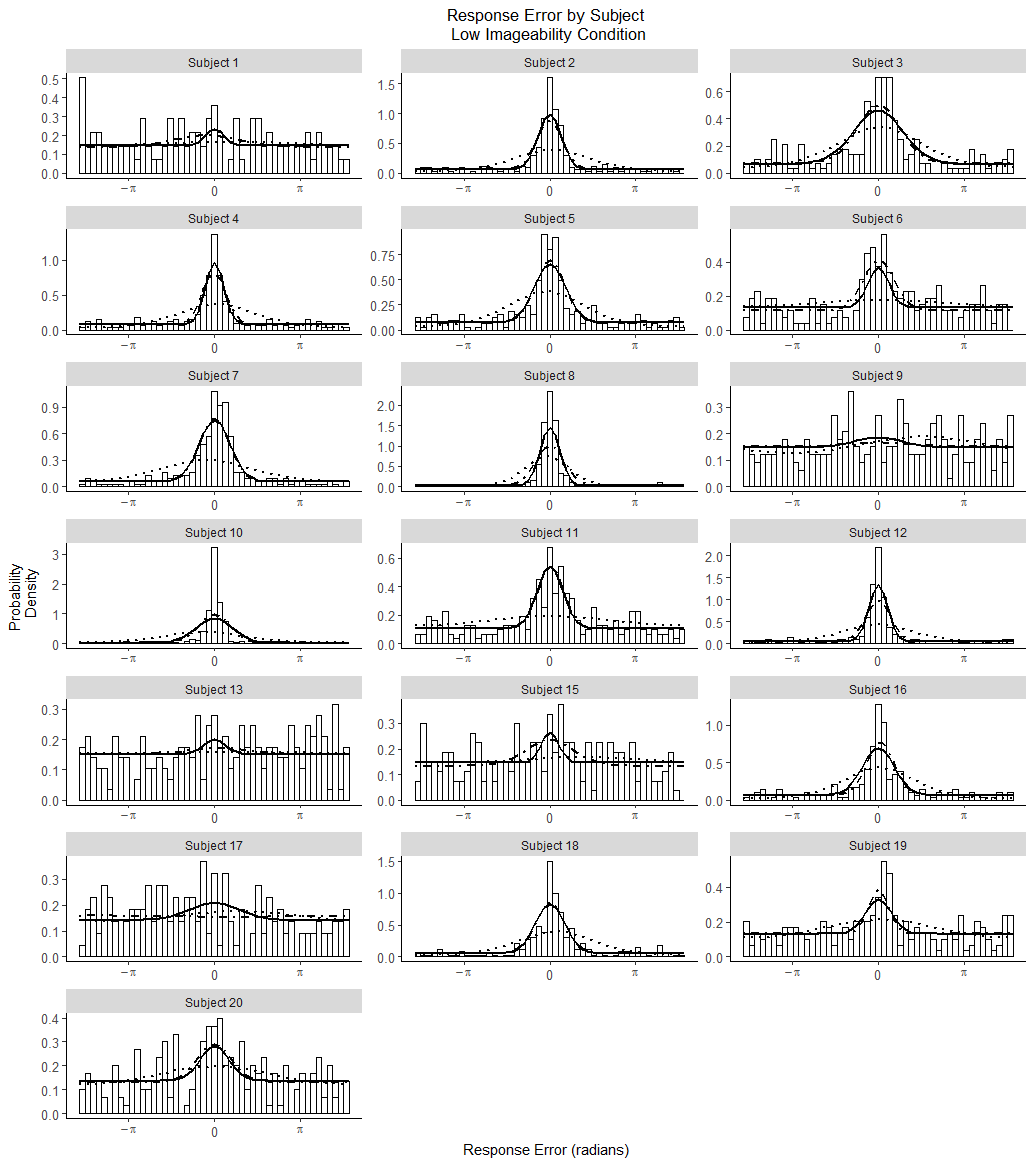
*Note.*  Precision refers to the precision of the information-driven retrieval process. *π* represents proportion of responses driven by information.

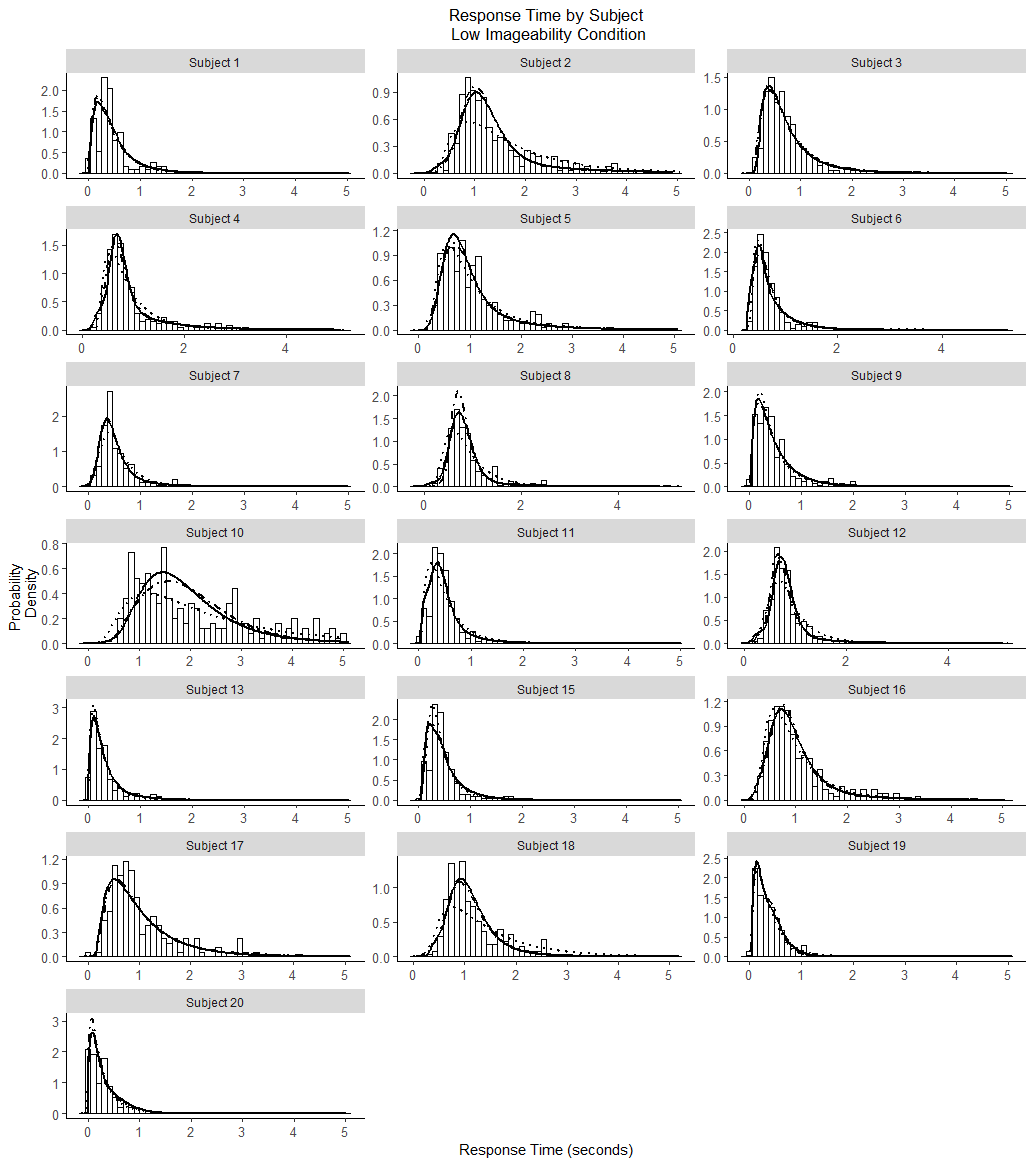
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Table X | | | | | |
| *Parameter Values for Best Fits of the Simple Mixture Model to Recognized Individual Data.* | | | | | |
|  | | Low Imageability | | High Imageability | |
| Participant | | Precision | *π* | Precision | *π* |
| High Accuracy |  |  |  |  |  |
|  | 2 | 24.51 | 0.66 | 17.36 | 0.73 |
|  | 3 | 11.76 | 0.45 | 10.47 | 0.54 |
|  | 4 | 32.44 | 0.48 | 44.47 | 0.54 |
|  | 5 | 18.51 | 0.47 | 16.80 | 0.65 |
|  | 6 | 12.90 | 0.27 | 8.86 | 0.23 |
|  | 7 | 12.17 | 0.65 | 9.11 | 0.69 |
|  | 8 | 37.43 | 0.87 | 42.93 | 0.89 |
|  | 9 | 0.57 | 0.29 | 0.10 | 1.00 |
|  | 10 | 49.40 | 0.87 | 49.08 | 0.87 |
|  | 11 | 7.90 | 0.35 | 13.10 | 0.39 |
|  | 12 | 36.72 | 0.73 | 46.67 | 0.69 |
|  | 15 | 5.39 | 0.13 | 85.49 | 0.04 |
|  | 16 | 15.53 | 0.55 | 11.01 | 0.64 |
|  | 17 | 0.27 | 0.64 | 5.81 | 0.10 |
|  | 18 | 9.77 | 0.69 | 27.61 | 0.59 |
|  | 19 | 14.27 | 0.19 | 5.45 | 0.32 |
|  | 20 | 20.61 | 0.15 | 10.15 | 0.25 |
|  |  |  |  |  |  |
| Low Accuracy | 1 | 157.24 | 0.05 | 250.00 | 0.02 |
|  | 13 | 45.64 | 0.03 | 250.00 | 0.02 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Table X | | | | | |
| *Parameter Values for Best Fits of the Simple Mixture Model to Highly Recognized Individual Data.* | | | | | |
|  | | Low Imageability | | High Imageability | |
| Participant | | Precision | *π* | Precision | *π* |
| High Accuracy |  |  |  |  |  |
|  | 2 | 22.47 | 0.69 | 16.61 | 0.76 |
|  | 3 | 12.11 | 0.48 | 9.97 | 0.58 |
|  | 4 | 33.96 | 0.50 | 43.79 | 0.58 |
|  | 5 | 19.47 | 0.47 | 16.12 | 0.67 |
|  | 6 | 13.73 | 0.31 | 13.30 | 0.24 |
|  | 7 | 11.58 | 0.69 | 8.90 | 0.74 |
|  | 8 | 36.73 | 0.88 | 42.74 | 0.89 |
|  | 9 | 0.69 | 0.27 | 0.14 | 1.00 |
|  | 10 | 49.03 | 0.87 | 47.47 | 0.89 |
|  | 11 | 8.27 | 0.39 | 12.60 | 0.39 |
|  | 12 | 38.12 | 0.76 | 45.42 | 0.73 |
|  | 15 | 1.45 | 0.38 | 22.78 | 0.08 |
|  | 16 | 15.53 | 0.55 | 10.80 | 0.64 |
|  | 17 | 0.90 | 0.20 | 5.27 | 0.11 |
|  | 18 | 9.95 | 0.71 | 25.90 | 0.60 |
|  | 19 | 14.27 | 0.19 | 5.45 | 0.32 |
|  | 20 | 18.29 | 0.17 | 9.52 | 0.26 |
|  |  |  |  |  |  |
| Low Accuracy | 1 | 192.59 | 0.05 | 0.22 | 1.00 |
|  | 13 | 3.69 | 0.10 | 250.00 | 0.02 |









References

Atkinson, R. C., & Juola, J. F. (1974). *Search and decision processes in recognition memory*. WH Freeman.

Banks, W. P. (2000). Recognition and source memory as multivariate decision processes. *Psychological Science*, *11*(4), 267-273.

Batchelder, W. H., & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological review*, *97*(4), 548.

Bell, R., Mieth, L., & Buchner, A. (2017). Emotional memory: No source memory without old–new recognition. *Emotion*, *17*(1), 120.

Bowers, J. S., & Schacter, D. L. (1990). Implicit memory and test awareness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(3), 404.

Dube, C., Starns, J. J., Rotello, C. M., & Ratcliff, R. (2012). Beyond ROC curvature: Strength effects and response time data support continuous-evidence models of recognition memory. *Journal of Memory and Language*, *67*(3), 389-406.

Harlow, I. M., & Donaldson, D. I. (2013). Source accuracy data reveal the thresholded nature of human episodic memory. *Psychonomic Bulletin & Review*, *20*(2), 318-325.

Hautus, M. J., Macmillan, N. A., & Rotello, C. B. (2008). Toward a complete decision model of item and source recognition. *Psychonomic Bulletin & Review*, *15*(5), 889-905.

Klauer, K. C., & Kellen, D. (2010). Toward a complete decision model of item and source recognition: A discrete-state approach. *Psychonomic Bulletin & Review*, *17*(4), 465-478.

Malejka, S., & Bröder, A. (2016). No source memory for unrecognized items when implicit feedback is avoided. *Memory & cognition*, *44*(1), 63-72.

Mickes, L., Wais, P. E., & Wixted, J. T. (2009). Recollection is a continuous process: Implications for dual-process theories of recognition memory. *Psychological science*, *20*(4), 509-515.

Onyper, S. V., Zhang, Y. X., & Howard, M. W. (2010). Some-or-none recollection: Evidence from item and source memory. *Journal of Experimental Psychology: General*, *139*(2), 341.

Osth, A. F., Bora, B., Dennis, S., & Heathcote, A. (2017). Diffusion vs. linear ballistic accumulation: Different models, different conclusions about the slope of the zROC in recognition memory. *Journal of Memory and Language*, *96*, 36-61.

Osth, A. F., Fox, J., McKague, M., Heathcote, A., & Dennis, S. (2018). The list strength effect in source memory: Data and a global matching model. *Journal of Memory and Language*, *103*, 91-113.

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

Ratcliff, R., & Starns, J. J. (2013). Modeling confidence judgments, response times, and multiple choices in decision making: Recognition memory and motion discrimination. *Psychological review*, *120*(3), 697.

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.

Rouder, J. N., Morey, R. D., Cowan, N., Zwilling, C. E., Morey, C. C., & Pratte, M. S. (2008). An assessment of fixed-capacity models of visual working memory. *Proceedings of the National Academy of Sciences*, *105*(16), 5975-5979.

Schurgin, M. W., Wixted, J. T., & Brady, T. F. (2018). Psychological scaling reveals a single parameter framework for visual working memory. *bioRxiv*, 325472.

Slotnick, S. D., & Dodson, C. S. (2005). Support for a continuous (single-process) model of recognition memory and source memory. *Memory & cognition*, *33*(1), 151-170.

Smith, P. L. (2016). Diffusion theory of decision making in continuous report. *Psychological Review*, *123*(4), 425.

Smith, P. L., & Ratcliff, R. (2009). An integrated theory of attention and decision making in visual signal detection. *Psychological review*, *116*(2), 283.

Smith, P. L., Saber, S., Corbett, E. A. & Lilburn, S. D. (2019). Modelling the speed and accuracy of continuous outcome colour decisions: Metric and categorical effects. Paper presented at the The Australian Mathematical Psychology Conference, Melbourne, Australia, February.. Abstract retrieved from: http://mathpsy.ch/abstracts/smith-26/

Starns, J. J., Hicks, J. L., Brown, N. L., & Martin, B. A. (2008). Source memory for unrecognized items: Predictions from multivariate signal detection theory. *Memory & Cognition*, *36*(1), 1-8.

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(6), 1341.

Yonelinas, A. P. (1999). The contribution of recollection and familiarity to recognition and source-memory judgments: A formal dual-process model and an analysis of receiver operating characterstics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(6), 1415.

Yonelinas, A. P., & Parks, C. M. (2007). Receiver operating characteristics (ROCs) in recognition memory: a review. *Psychological bulletin*, *133*(5), 800.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233.