Journal of Memory and Language 103 (2018) 91–113

Contents lists available at ScienceDirect

Journal of Memory and Language

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

The list strength effect in source memory: Data and a global matching model

Adam F. Osth[a,*], Julian Fox[a], Meredith McKague[a], Andrew Heathcote[b], Simon Dennis[a]

[a] The University of Melbourne, Australia
[b] The University of Tasmania, Australia

ARTICLE INFO

Keywords:
Recognition memory
Source memory
Global matching models
List strength effect

ABSTRACT

A critical constraint on models of item recognition comes from the list strength paradigm, in which a proportion of items are strengthened to observe the effect on the non-strengthened items. In item recognition, it has been widely established that increasing list strength does not impair performance, in that performance of a set of items is unaffected by the strength of the other items on the list. However, to date the effects of list strength manipulations have not been measured in the source memory task. We conducted three source memory experiments where items studied in two sources were presented in a pure weak list, where all items were presented once, and a mixed list, where half of the items in both sources were presented four times. Each experiment varied the nature of the testing format. In Experiment 1, in which each study list was only tested on one task (item recognition or source memory), a list strength effect was found in source memory while a null effect was found for item recognition. Experiments 2 and 3 showed robust null list strength effects when either the test phase (Experiment 2) or the analysis (Experiment 3) was restricted to recognized items. An extension of the Osth and Dennis (2015) model was able to account for the results in both tasks in all experiments by assuming that unrecognized items elicit guess responses in the source memory task and that there was low interference among the studied items. The results were also found to be consistent with a variant of the retrieving effectively from memory model (REM; Shiffrin & Steyvers, 1997) that uses ensemble representations.

Introduction

A distinction in episodic memory concerns the difference between information about learned content and the context in which it occurred. A common memory failure is when one remembers a fact or detail but has no memory for where he or she learned the information. The relationship between memory for content and context is studied in the laboratory using the item recognition and source memory paradigms. In the item recognition paradigm, participants study a list of items and at test are asked to discriminate between studied items (targets) and unstudied items (lures). The source memory paradigm presents participants with a set of items in different sources, such as different font colors, studied locations, or sensory modalities. At test, participants judge which source studied items were presented in.

A number of computational models of decision making have been developed to explain the relations between item and source memory (e.g.; Banks, 2000; Batchelder & Riefer, 1990; DeCarlo, 2003; Glanzer, Hilford, & Kim, 2004; Hautus, Macmillan, & Rotello, 2008; Klauer & Kellen, 2010; Slotnick & Dodson, 2005; Yonelinas, 1999). These models fall into several frameworks including multivariate signal detection theory, in which participants make decisions based on continuous latent strengths (SDT: Banks, 2000), discrete state models (Batchelder & Riefer, 1990; Klauer & Kellen, 2010), or a combination of continuous latent strengths and discrete states (Yonelinas, 1999).

While such models yield useful predictions about the shapes of the receiver operating characteristic (ROC) in each task (Slotnick & Dodson, 2005) and whether source memory is accurate without item memory (Starns, Hicks, Brown, & Martin, 2008), they are generally mute with respect to manipulations that often concern memory researchers, such as the effects of recency (Monsell, 1978), list length (Dennis, Lee, & Kinnell, 2008; Strong, 1912), list strength (Ratcliff, Clark, & Shiffrin, 1990), and word frequency (Glanzer & Adams, 1985), although, as addressed later, the Hautus et al. (2008) model makes one specific prediction with regard to the list strength paradigm in source memory. This is because these models define the form of the decision variable but are agnostic as to the encoding, storage, and retrieval assumptions that give rise to it. In contrast, the class of global matching models has made such specifications (Clark & Gronlund, 1996). In global matching models, memory strength is determined by the similarity between the retrieval cues and each stored item in memory; these similarities are summed (or averaged; Shiffrin & Steyvers, 1997) to produce a single strength value that can be compared to a response

---

[*] Corresponding author.
E-mail address: adamosth@gmail.com (A.F. Osth).

https://doi.org/10.1016/j.jml.2018.08.002
Received 16 January 2018; Received in revised form 31 July 2018

criterion to make a decision. Collectively, the current generation of episodic recognition models have been successful in explaining all of the aforementioned episodic memory phenomena in item recognition (e.g.; Dennis & Humphreys, 2001; Nosofsky, Little, Donkin, & Fific, 2011; Osth & Dennis, 2015; Shiffrin & Steyvers, 1997).

Nonetheless, many recent mechanistic models in the episodic memory literature have often been restricted to a single task and have rarely provided joint accounts of multiple memory tasks (but see Lehman & Malmberg, 2013, for a noteworthy exception). Hintzman (2011) criticized this tendency and argued that this has been leading to limited conclusions about the nature of memory as a whole. Consistent with this criticism, current mechanistic models of recognition memory have experienced little, if any, extension to the source memory paradigm. The current article attempts to fill this gap by testing one of the major constraints of episodic memory models, the list strength effect (LSE), in a source memory paradigm, and further introduce an extension of the Osth and Dennis (2015) model to provide a joint account of the results from both item recognition and source memory. The list strength paradigm asks the question *can strengthening a memory cause forgetting of other memories?*

The list strength paradigm: data and model predictions

A prediction of global matching models is that as the number of items in memory is increased, performance should decrease. In these models, each item in memory has variation in its similarity to the retrieval cues, so that as the number of items in memory is increased, the number of variance components that contributes to the decision increases and the signal-to-noise ratio is reduced. Ratcliff et al. (1990) found that the models yielded the same predictions for the case of repetitions of the list items; repetitions are treated in the same manner as increases in the number of studied items and contribute additional noise at retrieval.

To understand how this prediction is manifested, consider a sequence of study list items such as ABCD. Most models would predict that strengthening A and B via study time and/or repetition should increase performance on A and B. The counter-intuitive prediction that emerged from these models is that strengthening A and B should *impair* performance on C and D. This prediction can be tested by comparing lists with different compositions of strengthened and non-strengthened items, such as a pure weak list where all items are presented once (ABCD) and a mixed list where half the items are presented once and half the items are presented four times (AAAABBBBCD). The original global matching models predicted that performance of the once presented items (C and D) should be worse in the mixed list than in the pure weak list due to the extra interference from the repetitions of A and B. The list with more repeated items would be considered a list with higher *list strength*.

A large number of experiments tested and disconfirmed this prediction: increasing the strength of a set of studied items does not impair performance of the other items on the list for the case of item recognition with word stimuli (Diana & Reder, 2005; Hirshman, 1995; Kahana, Rizzuto, & Schneider, 2005; Ratcliff et al., 1990; Ratcliff, McKoon, & Tindall, 1994; Ratcliff, Sheu, & Gronlund, 1992; Shiffrin, Huber, & Marinelli, 1995; Yonelinas, Hockley, & Murdock, 1992) although small effects of list strength have been found with non-word stimuli such as faces and fractals (Norman, Tepe, Nyhus, & Curran, 2008; Osth, Dennis, & Kinnell, 2014). One should note that the free recall task contrasts with recognition memory in that increasing list strength has been shown to substantially impair performance when strengthening is achieved via spaced presentations (Tulving & Hastie, 1972) but not when strengthening is achieved via massed presentations or depth of processing (Malmberg & Shiffrin, 2005). As a consequence of this failed prediction, amendments to the global matching framework were proposed that enabled the models to predict a null list strength effect (LSE) in item recognition. One such modification was the differentiation hypothesis, in which repetitions accumulate into a single strong memory trace that is more responsive to its own cue but less responsive to other cues (Shiffrin, Ratcliff, & Clark, 1990). The latter component implies that strong memory traces generate less interference than weak traces, whereas in the older models the opposite was the case.

Another class of models has argued that the null LSE is more indicative of interference stemming from sources other than the studied items (Dennis & Humphreys, 2001; Murdock & Kahana, 1993a, 1993b; Osth & Dennis, 2015). While initial models assumed that memory is a “blank slate” before presentation of the study list,[1] these models instead assume an interference contribution from pre-experimentally learned memories consisting of prior occurrences of the cue word (context noise) or from other memories in general (background noise). When such interference contributions are substantial, interference from the additional repetitions in a list strength paradigm produces only a negligible increase in overall interference, allowing the models to predict null effects of list strength.

To our knowledge, none of these models which have been successful in addressing benchmark phenomena in item recognition have been applied to the source memory task. A simple extension of these models to source memory would involve binding each item to its source at study; at test the probe item would be cued with each of the studied sources and the memory strengths of each source cue would be compared. An example is depicted in Fig. 1, where “truck” and “joker” were studied in source A (red) and “sky” and “phone” in source B (green). At test, when prompted with a cue such as “truck”, in order to make a judgment as to which source “truck” was studied in participants could cue memory with a binding of “truck” in source A and match it to the contents of memory to obtain the memory strength for source A (s_A). Subsequently (or in parallel), the participant could cue memory with a binding of “truck” in source B and match it to the contents of memory to obtain a memory strength for source B (s_B). The difference between the memory strengths for source A and B could be used to make a decision - if this difference exceeds a decision criterion (ϕ_{source}) source A would be chosen, otherwise source B would be chosen.

Although this mechanism is similar to item recognition, the representational structure of the memory set in the source memory task can lead to different predictions. In item recognition, a word such as “truck” receives its strongest contribution from its own representation in memory, while the other items on the list produce much smaller degrees of match, due to the fact that they bear little resemblance to the retrieval cue. However, in source memory, half of the items in the list match the source cue, meaning that source memory can resemble cases where half of the representations in memory bear a high similarity to the retrieval cues.

We found this higher similarity in the task was sufficient to induce an LSE in the original version of the retrieving effectively from memory model (REM: Shiffrin & Steyvers, 1997); these simulations are detailed later in the General Discussion. This was somewhat surprising because in REM strengthening items produces differentiation of the memory traces, which should reduce the interference contribution from strong memory traces and produce a null LSE. However, differentiation only reduces interference when the similarity between the trace and the cue is relatively low. When the similarity is high, which is the case when 50% of traces match the source cue, interference increases with strength (Criss, 2006). However, later formulations of REM allow additional ensemble features that are unique to a binding between items or features (Criss & Shiffrin, 2005). Interestingly, ensemble features

---

[1] A reviewer pointed out that the central commitment of such models is not that there are no memories, but that any interference from prior memories is negligible. Another possibility is that interference from prior memories is eliminated because the context of the study list is sufficiently isolated from prior memories.