Journal of Memory and Language 66 (2012) 109–122

Contents lists available at SciVerse ScienceDirect

Journal of Memory and Language

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

Associative information in memory: Evidence from cued recall

William R. Aue *, Amy H. Criss, Nicholas W. Fischetti

Department of Psychology, Syracuse University, Syracuse, NY 13244, United States

ARTICLE INFO

Article history:
Received 5 February 2011
revision received 10 August 2011
Available online 14 September 2011

Keywords:
Cued recall
Episodic memory
Memory models
List discrimination
Associative memory

ABSTRACT

The representation of item and associative information in episodic memory was investigated using cued recall and single item recognition. In the first four experiments, participants studied two lists constructed such that some items presented in a pair during List 1 were rearranged to create new pairs in List 2 and were accompanied by pairs exclusive to List 2. List 1 was composed of either word–face pairs (Experiment 1 and 3) or word–word and face–face pairs (Experiment 2 and 4). Participants were tested for their memory of the second list containing only word-face pairs. When the test required associative information (i.e., cued recall), interference was specific to pair-type. Specifically, repeating items in the same pair-type across lists led to a greater number of correct and incorrect responses but repeating items in different pair-types did not change performance. When the test did not require associative information (i.e., single item recognition), interference was not specific to pair-type; hits and false alarms were higher for items presented on List 1. In the final experiment, the number of studied pairs and single items was independently manipulated and a pair-specific list length effect was observed. That is, performance in cued recall was modulated by the number of studied pairs, but not the number of studied individual items. Together, these data suggest that information specific to the pair-type (i.e., associative information) is stored and strategically utilized during memory search.

Introduction

Access to associative information plays a substantial role in everyday memory functioning. From helping to put a name with a face to understanding the meaning of a word based on its context, the information that binds independent pieces of information is critical to many facets of memory. However, how this associative information is represented in memory and how it differs, if at all, from item information is not fully understood. Underwood (1969) postulated that many different types of attributes, or features, may be stored for any event. For instance, item and context information are attributes commonly hypothesized to be stored during study. *Item features* include orthographic, visual, semantic, and other information that defines the item and *context features* include information such as temporal, spatial, and modality information describing the circumstances in which the item was encountered. The focus of the current paper is the nature of *associative features* which includes information indicating that a *pair* of items was encountered together.

In the laboratory, episodic memory for associations is measured by presenting pairs (e.g., A–B, C–D) at study and asking participants to remember that the two items appeared together. Two common approaches for measuring memory for associations are associative recognition and cued recall. During an associative recognition task, participants are tested on their ability to discriminate between intact pairs and rearranged pairs. Intact pairs are pairs that are an identical match to a studied pair (e.g., A–B), and rearranged pairs are composed of items from two different studied pairs (e.g., A–D). Thus, the participant

* Corresponding author. Address: Syracuse University, Department of Psychology, 430 Huntington Hall, Syracuse, NY 13244, United States. Fax: +1 315 443 4085.

E-mail addresses: william.aue@gmail.com (W.R. Aue), amy.criss@gmail.com (A.H. Criss).

0749-596X/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2011.08.002

must decide whether the pair of items was studied together in the earlier list. This is where a distinction between item and associative information becomes critical. Both items in a rearranged pair have been studied, thus participants cannot rely on their memory for the individual items A and D (i.e., item information). To be successful at this task, they must rely on their memory for which items occurred together (i.e., associative information). During a cued recall test, participants are provided with one member of the studied pair (e.g., A) and asked to generate its partner (e.g., B). Again, memory for an individual item alone is insufficient to be successful. Participants must remember B, but must also remember that B was studied with a specific item (i.e., A). Thus the critical difference between cued recall and associative recognition is that in cued recall, participants must derive the necessary associative information using only a single-item probe. In this sense, cued recall is similar to single item recognition because memory is cued with a single item. How this associative information is represented in memory and how it is used during the retrieval process is poorly understood.

The goal of this paper is to better describe the nature of associative features using cued recall. First, we describe a few of the many different theoretical approaches for representing associative features. *Emergent feature models* (e.g., Eich, 1982; Murdock, 1982, 1997; Murnane & Shiffrin, 1991) suggest that something unique about the pair is generated and stored in memory.[1] This emergent information exceeds the information provided by the items alone. Such models represent associative information as a unique set of features distinct from the individual item features. Emergent features are assumed to be an additional set of features that are independent of item information and may be stored better or worse than item features, or not at all, depending on the goals at encoding. *Associative link models* (e.g., Diana, Reder, Arndt, & Park, 2006; Gillund & Shiffrin, 1984; Reder et al., 2000) represent an association as a connection between two separate items. Similar to the emergent features approach, the link may or may not be created during encoding depending on the goals of the learner and is specific to the two items. However unlike emergent features, it does not contain new information beyond that provided by the individual items. Unlike the previous models, *co-occurrence models* do not provide a distinct representation of associative information. Co-occurrence models represent associations as the simple co-occurrence of two items (e.g., Hintzman, 1984, 1988; McClelland & Chappell, 1998; Shiffrin & Steyvers, 1997). Thus, no associative information is stored per se; rather the association is represented by the fact the features of each item are stored together within the same memory trace or sharing the same context. Emergent features and associative links both allow for differential encoding of item and associative information. Items may be stored well while associative information is poorly stored (or vice versa). This is not the case for co-occurrence models where items and associations are composed of the same information. If items are stored with high accuracy, then so too is the association.

Finally, consider *class attributes* (Criss & Shiffrin, 2004, 2005; Underwood, 1969). Class attributes provide information about the nature of items or pair. Underwood (1969) assumed that people use class attributes to selectively access subsets of items within memory. For example, he pointed out that if the task is to recall digits, people do not recall words. If a person is trying to name a tool, she does not generate the name of a food item (see also the selector mechanism of Underwood & Schulz, 1960). In other words, humans seem able to limit the search of memory to a subset based on type. Criss and Shiffrin (2004, 2005) empirically studied class attributes by manipulating the type of pair (e.g., a word–face pair vs. word–word pair) that items were studied in as an associative class attribute. Note that class attributes could be used during a memory search in conjunction with the other types of associative features (emergent, co-occurrence, or links).

The balance of evidence supports the idea that unique associative features are stored in addition to item features (e.g., Gronlund & Ratcliff, 1989; Hockley & Consoli, 1999; Hockley & Cristi, 1996a, 1996b; Kelley & Wixted, 2001; Nobel & Shiffrin, 2001; see Clark and Gronlund (1996) for a review). For example, Hockley and Cristi (1996a) had participants study pairs of items under one of two instructional conditions. Participants were told to either focus on remembering the individual items, or to focus on remembering the episodic relationship between the items in a pair. Their memory was then tested using either associative recognition or single item recognition. On a test of associative recognition, Hockley and Cristi found that participants who focused on item information during study did worse than those who focused on associative information. Moreover, there was no difference in single item recognition performance between the two study conditions. This is consistent with separate representations for items and associations, representations that are stored independent of one another (e.g., emergent features or associative links).

While the aforementioned data provided evidence of dissociation between item and associative information, Criss and Shiffrin (2004) provided an experimental framework within which to better understand the nature of the associative information. They had participants study single lists of intermixed word–face (WF), word–word (WW), and/or face–face (FF) pairs to take advantage of Underwood's (1969) notion of *class attributes*, features that represent the information about the class of items. The relative number of each pair-type was manipulated. In tests of associative recognition, Criss and Shiffrin (2004) observed pair-type-specific list-length effects. For example, adding WF pairs to a study list harmed accuracy for WF pairs but not WW or FF pairs. In contrast, in tests of single item recognition, they found no effect of the relative number of each pair-type; instead performance was determined by the total list length. Based on their data, Criss and Shiffrin (2004) argued for two types of associative features: class attributes that identify the type of pair stored and emergent features that represent unique information about the combination of the two items.

[1] We use the term emergent features, others have used terms such as higher order units (Shiffrin, Murnane, Gronlund, & Roth, 1989), ensembles (Murnane, Phelps, & Malmberg, 1999), convolutions (Eich, 1982; Murdock, 1982), etc.

Criss and Shiffrin (2005) further investigated the nature of stored associative information using an A–B, A–D interference design.[2] Participants studied two lists. The second study list consisted of only WF pairs. The composition of the first list was either WF pairs or WW and FF pairs. Some items appeared on both lists but in different pair combinations (e.g., AB, CD, EF studied on list 1 and AD, CF studied on list 2) and some items appeared on just one list. The critical manipulation across experiments was whether repeated items appeared as part of the same or different type of pair across the two lists. The test was list discrimination (or equivalently, an exclusion task, c.f. Jacoby, 1991) and participants were asked to accept items (singles or pairs) from List 2 and reject everything else. In single item recognition tests of both faces and words, list composition did not matter, HRs were higher for repeated items and FARs to List 1 items were equally high regardless of whether List 1 contained the same or different type of pair as List 2. However the same was not true for associative recognition tests where targets were pairs from List 2 and foils were rearranged pairs composed of two items from the same condition (e.g., AF in the example above). Participants had higher HRs and FARs for pairs containing repeated items only when the items were repeated in the same type of pair across list. Studying individual items in the same pair-type on both study lists lead to a higher probability of calling a pair old relative to pairs composed of items studied just once. However, when repeated items appeared in different pair-types across lists the effect was eliminated. In summary, when making a memory decision about individual items pair-type plays no role, but when making a memory decision about a pair the decision is modulated by other members of that same pair-type. In other words, people flexibly use associative features in item vs. associative recognition tests.

Criss and Shiffrin (2005) fit the data using a version of the retrieving effectively from memory (REM; Shiffrin & Steyvers, 1997) model. The critical change in the model was that the memory vector now contained emergent features of the pair and class attribute information along with each set of item information. Most importantly for the current studies, Criss and Shiffrin's model used the emergent features of the test pair to probe memory. This is simply not possible for cued recall given that the memory probe is a single item. Criss (2005) suggested that the pattern of data is cue dependent and that a cued recall version of this paradigm should result in data that are more similar to single item recognition because both tasks use the same type of cue to probe memory.

The focus of the current paper was to investigate the use of associative features in cued recall within the paradigm developed by Criss and Shiffrin (2005). In their associative recognition experiments the cue used to probe memory was a pair, allowing the probe of memory to contain item information, emergent features about the pair, and class attribute features about the type of pair. In the cued recall task used in this paper, the cue used to probe memory is a single item and potentially class attributes indicating the type of pair under consideration. The constraints that cued recall places on probe information along with the Criss and Shiffrin (2005) design allows us to disentangle the information provided in the cue from the type of to-be-remembered information. Emergent features are not part of the probe in cued recall. However the class attribute or type of pair is an implicit part of the cue. If class attributes serve to limit the memory search in cued recall to a specific subset of items, then cued recall should be influenced by pairs of the type being cued and no other pairs (mimicking associative recognition performance). If instead the search of memory is limited by the item features in the cue, then cued recall should not be influenced by pair-type (mimicking single item recognition).

[2] The corresponding terminology from paired associate literature is A–Br.

Experiment 1

The purpose of this experiment was to investigate the use of associative features in cued recall. To do this we used the Criss and Shiffrin (2005) design, illustrated in Fig. 1. Participants studied two lists. The first list contained WF pairs. The second list also contained WF pairs, some created from rearranging items from List 1 pairs and others created with items unique to List 2. Using this design, Criss and Shiffrin (2005) found that in a test of associative recognition, participants had a higher HR and FAR to items that were repeated across list. They subsequently modeled the data by proposing that class attributes for pair-type and emergent features created from the unique combination of items were stored in addition to item and context features. Given that the probe in cued recall is a single item it is not possible to use emergent features to search memory. If class attributes are used to limit the search of memory in cued recall to the relevant subset (e.g., WF pairs), then interference from List 1 pairs should be present if List 1 contains the same type of pair as List 2 (Experiment 1), but absent when List 1 contains a different type of pair (e.g., WW and FF pairs, Experiment 2).

Method

Participants

Thirty-five participants from Syracuse University participated for a class requirement.

Materials

The face stimuli used were the 210 black and white photographs of faces described in Criss and Shiffrin (2004, 2005). Each face was standardized for head orientation, eye level, and nose position. The word stimuli used were 800 high frequency words ($M = 130.66$ from Kucera & Francis, 1967 or alternatively log frequency $M = 10.46$ in the Hyperspace Analog to Language corpus of Balota et al., 2007; Lund & Burgess, 1996).

Design and procedure

Participants completed the experiment individually on a Windows based computer running Authorware (v. 7.0). The experiment was a within-subject design with two conditions (rearranged and List 2 only). Participants studied 2 lists, followed by a test for the second list. Study was

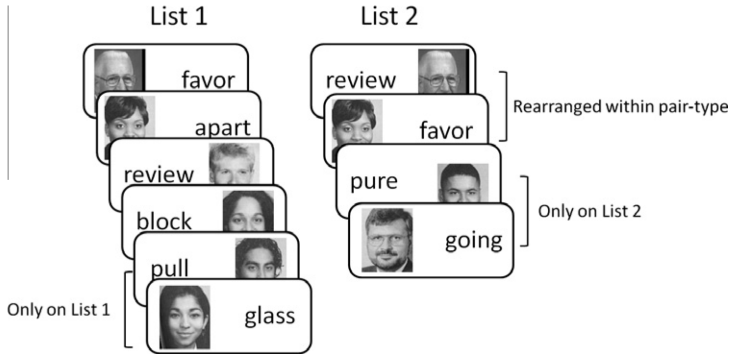

Fig. 1. The design employed in Experiments 1 and 3. Word–face pairs studied during the first study list were rearranged to create new word–face pairs for the second study list. The second study list also included pairs composed of items that were exclusive to the second study list. In Experiment 1 participants were then tested for their memory of the second list using cued recall. In Experiment 3 their memory for the second list was tested using single item recognition.

incidental to avoid instances where participants might chain together or co-rehearse pairs. List 1 and 2 each contained 24 WF pairs. As illustrated in Fig. 1, half of the List 2 pairs were created by rearranging items from List 1, called *rearranged* pairs. The other half of the pairs were created from items unique to List 2, called *List 2 only* pairs. (The remaining 12 pairs on List 1 were not tested here but were included as foils in Experiments 3 and 4.) The pairing, the re-pairing for rearranged pairs, and the assignment of pair to condition was randomly chosen for each participant.

Items in each pair were presented side-by-side on the screen for 3 s, with the left/right position of the face and word randomly chosen on each trial for each participant. Different encoding tasks were used across the two study lists to facilitate list discrimination. The encoding tasks employed are same as those successfully used in Criss (2005), Criss, Aue, and Smith (2011), and Criss and Shiffrin (2004, 2005). Moreover, Criss et al. (2011) found no difference in cued recall performance for pairs studied under one encoding task vs. another. For each List 1 pair, the participant was asked to rate the degree of association between the items on a 9-point scale (1 = *not at all associated* and 9 = *highly associated*). After the last pair of List 1 was rated, subjects read a comic for 60 s. For each List 2 pair, the participant were asked to generate a sentence for each pair and then rate the difficulty of doing so on a 9-point scale (1 = *very easy* and 9 = *very difficult*). After the second study list, participants completed a 60 s distracter task that involved keeping a running summation of 20 single digits (i.e., each digit was presented for 3 s). Immediately following the distracter task, subjects received each face from List 2, presented one at a time in a self-paced fashion. The order of faces was randomized anew for each subject. They were asked to type the word the face cue was paired with on the most recent list. Participants were given the option to respond that they did not recall the word the face was paired with by typing "no". The entire session lasted approximately 30 min.

All responses were coded as either correct, intrusion, or no recall. A correct response consisted of the target word or minor errors such as misspellings (e.g., braec instead of brace) or added suffixes (e.g., walks instead of walk). All other responses were coded as an intrusion. Trials when participants responded "no" or did not respond were coded as no recall.

Results and discussion

For all results a repeated measures analysis of variance (ANOVA) was performed separately for correct responses, intrusions, and no response trials unless otherwise stated. As can be seen in Fig. 2, participants had significantly more correct responses for rearranged pairs ($M = .312$, $SE = .025$) relative to the List 2 only pairs ($M = .224$, $SE = .028$; $F(1,34) = 9.79$, $p = .004$, $\eta_p^2 = .224$). Participants also made significantly more intrusions for rearranged items ($M = .236$, $SE = .026$) relative to List 2 only items ($M = .167$, $SE = .019$; $F(1,34) = 6.14$, $p = .018$, $\eta_p^2 = .153$). The increased number of correct responses and intrusions were accompanied by significantly fewer no recall responses ($M = .452$, $SE = .029$) for rearranged items relative to List 2 only ($M = .610$, $SE = .026$; $F(1,34) = 22.33$, $p < .001$, $\eta_p^2 = .396$). The higher levels of accuracy for the rearranged condition may, at first, appear to be a reversal of the standard proactive interference effect. However, this is misleading because intrusions for the rearranged condition are also higher. We discuss possible explanations for this finding in the General Discussion.

Although less critical to interpreting the primary results, we analyzed the different types of intrusions. Due to the limited number of intrusions, we could not analyze individual participants and instead considered only group data. Intrusions were coded as one of five types: if the intruded word appeared on both List 1 and 2 but was *not* the partner of the cue from List 1 it was coded as a List 1 and List 2 item, if the intruded word *was* the word studied with the cue on List 1 it was coded as a List 1 partner (this was only possible for rearranged pairs), if the intruded word was one that only appeared on List 2 it was coded as a List 2 only item, if the intruded word was one that only appeared on List 1 it was coded as a List 1 only item. Intrusions that did not fall into one of these categories were labeled as other and excluded from analyses ($N = 8$ for rearranged and $N = 16$ for List 2 only). A Pearson's

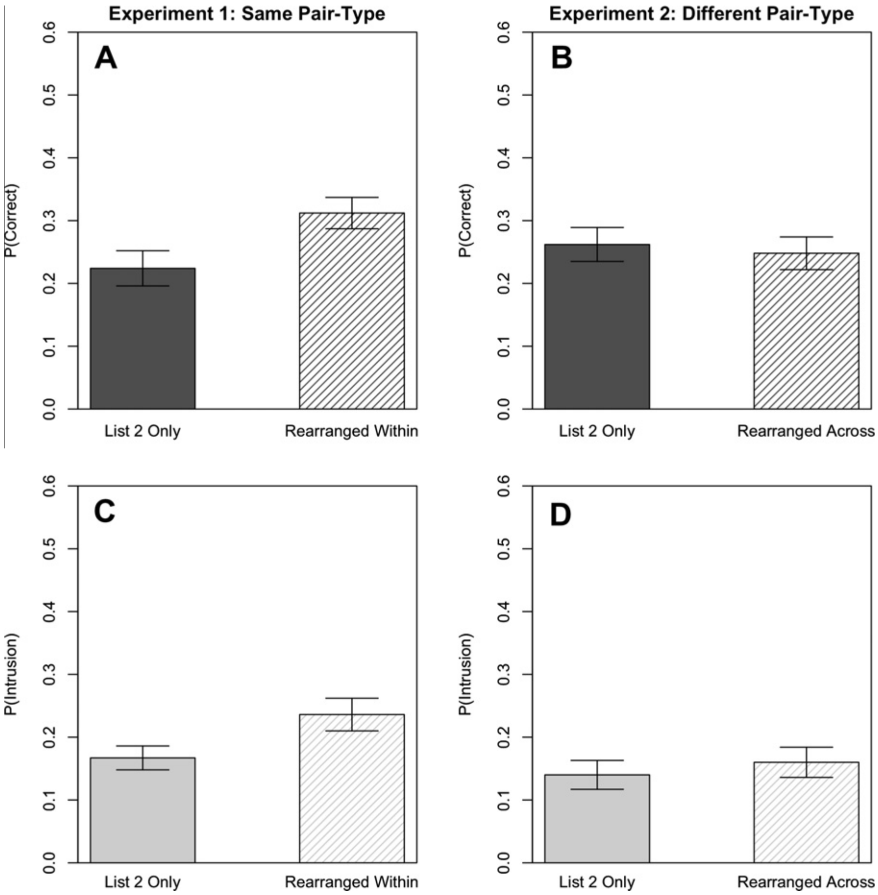

Fig. 2. The probability of providing a correct response (Panels A and B) or an incorrect response (Panels C and D) by condition for Experiments 1 and 2. In Experiment 1, both lists contain the same pair-type and items are rearranged into the same type of pair on List 2. In Experiment 1, there are significantly more correct and incorrect responses for Rearranged than List 2 only pairs. In Experiment 2, the study lists contain different types of pairs and items are rearranged into a different type of pair on List 2. In Experiment 2, there is no difference between conditions in either correct or incorrect responses.

Table 1
The number of each intrusion type by condition for Experiment 1.

	Experiment 1					
	Rearranged within			List 2 only		
Intrusion type	Possible	Expected	Observed*	Possible	Expected	Observed
List 1 only item	12	31.2	8	12	18.51	12
List 2 only item	12	31.2	14	11	16.98	16
List 1 partner	1	2.6	36	NA		
List 1 and List 2 item	10	26	33	12	18.51	26
Total	35	91	91	35	54	54

Note. The expected values for intrusions were generated based the number of *unique* intrusions possible for each condition. For example, for Rearranged pairs there were 12 non-target words that were studied only on List 1, 12 that were only on List 2, 10 that were on both lists, and 1 that was the List 1 partner of the cue. The number of possible intrusions was used to generate probabilities for each intrusion type which were then applied to the total number of intrusions for a given condition.

* $p < .05$.

chi-squared test was used to examine intrusion frequencies. The expected values for the analysis were derived from the proportion of intrusions possible for a given type, as shown in Table 1. Consider the rearranged condition. For any rearranged cue, there were 12 possible List 1 only words that could intrude, 12 List 2 only words, 10 words that were on both List 1 and List 2, and 1 List 1 partner. There were 91 total intrusions in the rearranged condition and 34% of them (12/35) are expected to be List 1 only words. Thus, the expected number of List 1 only intrusions was 31. Expected values for each intrusion type for each condition were calculated likewise.

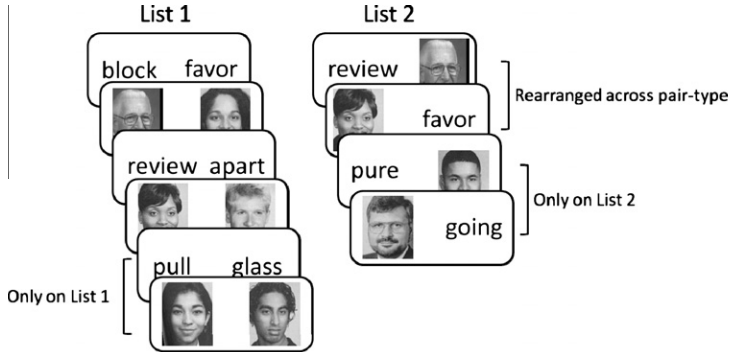

Fig. 3. The design employed in Experiments 2 and 4. Word–word or face–face pairs studied during the first study list were rearranged across pair-type to create new word–face pairs for the second study list. The second study list also included pairs composed of items that were exclusive to the second study list. In Experiment 2 participants were then tested for their memory of the second list using cued recall. In Experiment 4 memory for the second list was tested using single item recognition.

As can be seen in Table 1, intrusions for Rearranged pairs differed significantly from expected values (χ^2 (3, $N = 91) = 50.00$, $p < .001$). More intrusions tended to come from the List 1 partner of the cue and fewer intrusions tended to come from List 1 only or List 2 only (e.g., once presented) items. However, the intrusions for List 2 only pairs did not differ significantly from the expected values (χ^2 (2, $N = 54) = 2.68$, $p = .262$).

Using a similar design, Criss and Shiffrin (2005) and Overman and Becker (2009) showed a higher HR and FAR for rearranged items relative to List 2 only items in associative recognition. A similar pattern was observed here in cued recall – higher correct and incorrect responses to pairs that were rearranged across lists compared to pairs that were composed of items studied in just List 2. Emergent features are not present in the cued recall memory probe, which consists of a single face in this experiment. Thus this study presents some evidence that emergent features are not the only source driving the pattern of data observed here and in Criss and Shiffrin (2005), that is, interference from pairs of the same type. Two possibilities remain: either all items that match the cue contribute to memory performance, or class attributes restrict the search of memory to pairs of the same type. In the next experiment we evaluate these possibilities by creating List 1 from pairs of a different type (i.e., WW and FF) than List 2 (i.e., WF). If all item features similar to the memory cue contribute to performance, then we should observe data similar to that of Experiment 1 (increased response rate to rearranged vs. list 2 only pairs). If, however, only pairs of the same type contribute to performance, as is the case in associative recognition, then we should observe a different pattern of data than Experiment 1. Specifically, the influence of List 1 on cued recall performance should be minimal.

Experiment 2

The second experiment was identical to Experiment 1 with the exception of the pair composition in List 1, see Fig. 3. List 1 consisted of word–word (WW) and face–face (FF) pairs instead of word–face (WF) pairs. Thus for the rearranged condition, items were repeated in different pair-types across lists. If participants are able to use a pair-type class attribute to strategically search memory then the interference from List 1 should be minimal, in contrast to the findings of Experiment 1.

Method

Participants

Thirty-five participants from Syracuse University participated for a class requirement.

Materials

The materials used were the same as those described in Experiment 1.

Design and procedure

The details were the same as described for Experiment 1 with the critical exception of the List 1 pair-type. In the current experiment, List 1 was comprised of 12 WW pairs and 12 FF pairs.

Results and discussion

As can be seen in Fig. 2, there was no difference between the pairs composed of repeated items and the List 2 only pairs for correct recall ($F < 1$), intrusions ($F < 1$), or trials with no response ($F < 1$).

The results from Experiment 2 displayed a similar pattern as the Criss and Shiffrin (2005) associative recognition data: there was no interference from List 1 items when it contained a different pair-type than List 2. Criss and Shiffrin modeled their data by assuming that participants used class attributes to selectively access the relevant subset of memory (e.g., WF pairs), effectively isolating the second study list, thereby eliminating interference from List 1. The current data extend this to cued recall and suggest that participants use class attributes to search memory even when the probe consists of a single item (not a pair).

Intrusions were analyzed using the same method described in Experiment 1. For both conditions, 10 intrusions were excluded from the analysis because they were extra-experimental. The frequencies of intrusion types for both Rearranged pairs (χ^2 (2, $N = 57) = 15.42$, $p < .001$) and List 2 only pairs (χ^2 (2, $N = 46) = 21.61$, $p < .001$) differed signif-

Table 2
The number of each intrusion type by condition for Experiment 2.

	Experiment 2					
	Rearranged across			List 2 only		
Intrusion type	Possible	Expected	Observed*	Possible	Expected	Observed*
List 1 only item	12	19.54	3	12	16.8	0
List 2 only item	12	19.54	25	11	15.4	17
List 1 partner	NA			NA		
List 1 and List 2 item	11	17.92	29	12	16.8	32
Total	35	57	57	35	49	49

* $p < .05$.

icantly from the expected values. As can be seen in Table 2, this is driven by the lack of List 1 only intrusions and greater than expected List 1 and List 2 item intrusions. The lack of List 1 only intrusions is consistent with the idea that class attributes are used to effectively limit the search of memory to List 2 (e.g., WF pairs).

Lastly, the data for Experiments 1 and 2 were combined into a mixed ANOVA for correct responses and intrusions with condition (rearranged, list 2 only) as the within subject factor and experiment as the between subject factor. As is evident in Fig. 2, there was a significant interaction of condition and experiment for correct responses ($F(1,68) = 6.58$, $p = .013$, $\eta_p^2 = .088$). The interaction for intrusions approaches significance ($F(1,68) = 1.63$, $p = .206$, $\eta_p^2 = .023$).

In summary, Experiments 1 and 2 provide evidence that distinct associative features are stored in memory. This is evidenced by the interference when items were rearranged within pair-type, but not when items were rearranged across pair-type. The current data also suggest that pair-type specific interference is the result of flexibly engaging class attributes to access a subset of memory for associative memory tasks.[3] To be sure that this is the result of using class attributes, we must also show the absence of pair-type effects when pair specific class attributes are not relevant. Experiments 3 and 4 are the single item recognition analogs of Experiments 1 and 2 respectively. We expect to find interference from List 1 regardless of the type of pair(s) on that list.

Experiment 3

In Experiments 1 and 2, the effect of rearranging items within pair-type vs. across pair-type was limited to tasks requiring associative information. In both cued recall and associative recognition, performance is determined by pairs of the same type, but not affected by pairs of a different type. We now employ the same basic paradigm but test single item recognition to determine whether participants use class attributes in single item recognition as they do in cued recall. The cue presented to the participant is identical in single item recognition and cued recall (e.g., a single item). The difference between the tasks is the response required. In cued recall participants must generate the item studied with the cue but in single item recognition participants determine whether the cue had been studied on the most recent list. In other words, cued recall requires associative information and single item recognition does not. If class attributes defining the pair are used in single item recognition, then we should see the same pattern of pair-type specific interference as we do in cued recall. Note that class attributes may define the item-type, either a word or face. However, in this design an equal number of faces and words are studied in both conditions. The only thing that differs is how the items are combined into pairs. Experiments 3 and 4 are replications of the single item recognition experiments in Criss and Shiffrin (2005) but with details that match the cued recall studies presented here. Criss and Shiffrin found no evidence for the use of class attributes in single item recognition and we expect to replicate their results.

Method

Participants

Thirty-five participants from Syracuse University participated for a class requirement.

Materials

The materials used were the same as those described in Experiment 1.

Procedure

The study procedure and list composition was the same as described in Experiment 1 with a few exceptions. First, the length of the study lists was increased to 64 pairs to avoid possible ceiling effects. During the second study list, 32 of these pairs were rearranged across both lists and 32 were exclusive to the second study list. Second, participants were tested using single item recognition instead of cued recall. Testing consisted of showing participants 64 items, equally divided words and faces and targets and foils. Only one randomly chosen item from a given pair was tested. Of the targets, eight faces and eight words appeared in both List 1 and List 2 (in different pairs on each

[3] A reviewer pointed out that alternative explanations are possible. For example, perhaps FF and WW pairs interfere with WF pairs, but to a lesser extent (so minimal that it cannot be measured here or in Criss and Shiffrin (2004, 2005)). Perhaps the result is due to mixed lists (i.e., FF and WW pair-types) in Experiment 2 vs. pure lists (i.e., WF only) in Experiment 1 or practice with WF pairs in list 1 for Experiment 1 but not Experiment 2. Based on the results of Experiments 1 and 2, we cannot rule out these possibilities. However, we note that Criss and Shiffrin (2004) used mixed designs in associative recognition and found the same result: pair-type specific interference. Further, we have unpublished data using mixed lists in cued recall and find the same pattern of data.

list) and eight faces and eight words appeared just on List 2. Of the foils, eight faces and eight words were from List 1 and eight faces and eight words were novel to the experiment. Participants were shown the test item in the middle of the screen and asked if the item appeared in the most recent study list. Participants were provided with "Yes" and "No" response buttons. All other aspects of the experiment were identical to Experiment 1.

Results and discussion

A 2 × 2 repeated measures ANOVA comparing stimulus type (word vs. face) and pair-type (List 1 and 2 vs. List 2 only) was performed for both HRs and FARs. For HRs, there was a significant main effect of pair-type for HR. Participants had a higher HR for targets that appeared on both lists than to targets that only appeared on List 2 only ($F(1,42) = 59.85$, $p < .001$, $\eta_p^2 = .588$). There was no difference in HR for faces and words ($F < 1$). A similar pattern of data was observed for FAR. There was a significant main effect of pair-type for FAR, participants had a higher FAR for foils that appeared on the first list than to foils that only appeared during test ($F(1,42) = 81.5$, $p < .001$, $\eta_p^2 = .660$). There was no difference in FAR for faces and words ($F < 1$). There were no interactions between stimulus type and pair-type for FARs ($F < 1$) or HRs ($F < 1$). The means and standard error for pair-type, collapsed over stimulus type are plotted in Fig. 4. For archival purposes, data by stimulus type are presented in Table 3. This pattern of data replicates the basic pattern observed in Criss and Shiffrin (2005) of higher HR and FAR for repeated items. Note that this is interference, not a benefit, despite the increase in HR.

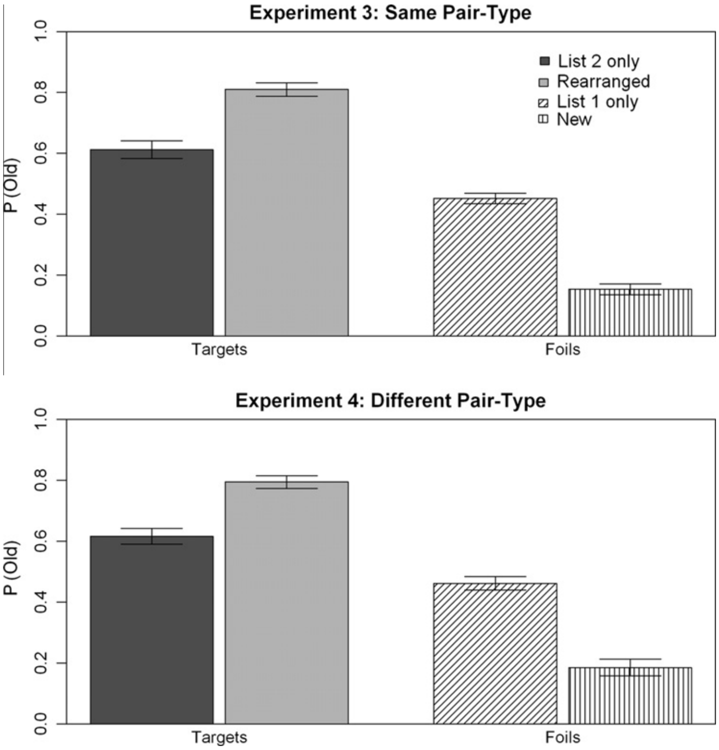

Fig. 4. The probability of endorsing an item as old in single item recognition for Experiments 3 and 4. Items that were studied more than once (i.e., Rearranged) were more likely to be correctly identified as old than items studied only once (i.e., List 2 only). Foils that had been studied previously on List 1 were more likely to be endorsed as old than novel foils. Critically, the same pattern of data holds regardless of whether the pairs are rearranged into the same type of pair (Experiment 3) or different type of pair (Experiment 4) across study lists.

Table 3
The probability of an "Old" response for each stimulus type in Experiments 3 and 4.

	Experiment 3				Experiment 4			
	Words		Faces		Words		Faces	
	M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Targets (Hits)								
Rearranged	.814	.025	.805	.030	.810	.027	.780	.023
List 2 only	.605	.023	.619	.026	.640	.030	.599	.033
Foils (FAs)								
List 1	.456	.034	.448	.035	.462	.030	.462	.028
New	.160	.024	.148	.019	.192	.035	.180	.026

Experiment 4

Experiment 4 is the single item recognition analog to Experiment 2. Study List 1 consisted of WW and FF pairs and List 2 consisted of WF pairs. Like Criss and Shiffrin (2005), we expect to find that participants do not use class attributes when the test only necessitates item information.

Method

Participants

Forty-three participants from Syracuse University participated for a class requirement.

Materials

The materials used were the same as those described in Experiment 1.

Design and procedure

The study design and procedure was identical to Experiment 3 except that the items on List 1 were composed of WW and FF pairs (32 each).

Results and discussion

A 2 × 2 repeated measures ANOVA comparing stimulus type and pair-type was performed for both HRs and FARs. For HR, there was a significant main effect of pair-type for HR. Participants had a higher HR for targets that appeared on both lists than to targets that only appeared on List 2 ($F(1,42) = 46.62$, $p < .001$, $\eta_p^2 = .526$). There was no difference in HR for faces and words ($F(1,42) = 1.19$, $p = .28$). A similar pattern of data was observed for FAR; there was a significant main effect of pair-type for FAR. Participants had a higher FAR for foils that appeared on the first list than to foils that only appeared during test ($F(1,42) = 110.72$, $p < .001$, $\eta_p^2 = .725$). There was no difference in FAR for faces and words ($F < 1$). There were no interactions between List 1 pair-type and stimulus type for HRs ($F < 1$) or FARs ($F < 1$). The means and standard error for pair-type, collapsed over stimulus type are plotted in Fig. 4. For archival purposes, data by stimulus type are presented in Table 3. As expected, the current data replicate both the results from Experiment 3 and the data observed in Criss and Shiffrin (2005).

Lastly, the data from Experiments 3 and 4 were entered into a 2 × 2 mixed ANOVA comparing stimulus type (word, face) and condition (rearranged, list 2 only) for both HRs and FARs across the two experiments. For HR, only the main effect of condition was significant ($F(1,84) = 105.94$, $p < .001$) with items from Rearranged pairs being better recognized than items from List 2 only pairs. Also, for FAR only the main effect of condition was significant ($F(1,84) = 185.37$, $p < .001$) with items from List 1 having a higher FAR than items from New to the test list.

Based on the data from Experiments 3 and 4 we suggest that associative information is not used when memory for individual items is evaluated. The advantage of rearranged items over List 2 items in both cases can be explained by the fact that rearranged items were experienced twice over the course of the experiment and as such would be more familiar than items experienced only once. Likewise, for foils, the items that were experienced in List 1 would be more familiar than items that were unique to the test list. Critically, there is no evidence that participants use associative features at test. During test participants are instructed to decide whether the item appeared on the most recent list. This implies that the item was a member of a WF pair, nevertheless participants do not use class attributes to aid their memory search. In the single item experiments, cue-type (words and faces) was intermixed but in the cued recall experiments, cue-type (face) was constant. We note that earlier studies of associative recognition also intermixed cue-type and found pair-type-specific interference, thus we do not believe that mixed vs. pure cue type is responsible for the difference patterns of interference in single item recognition and cued recall (Criss, 2005; Criss & Shiffrin, 2004). Note that the encoding task, instructions, and the incidental nature of the task were identical for the cued recall (Experiments 1 and 2) and single item recognition (Experiments 3 and 4) versions. The differences only arise when the test begins, thus the processes driving use or disuse of associative information must occur at retrieval. Participants appear to reserve the use of associative features for associative-based memory tasks.

Experiment 5

Thus far we have demonstrated that participants are able to selectively access subsets of items within memory through the use of class attributes and do so for tests of associative memory but not item memory. The purpose of Experiment 5 was to provide converging evidence for the use of class attributes in cued recall by using a design that allowed for searching within a single list rather than across lists. In an associative recognition experiment, Criss and Shiffrin (2004) manipulated the number of WW, FF, and WF pairs on a study list and then tested participants on associative recognition of the pairs. They found that adding pairs to a study list only harmed performance for pairs of the same type but did not affect performance for other pair-types. For example, adding WW pairs only harmed performance for WW pairs while leaving performance for WF pairs unchanged. This occurs despite the fact that WW, FF, and WF pairs share the same class of constituent items. In the current experiment participants studied a single list and list length was manipulated by adding pairs of the same or different type. Participants were then tested using cued recall. If participants use class attributes to limit memory search, then adding pairs of the type that is being tested should harm performance but adding pairs of a different type should not.

Method

Participants

Two hundred-three students from Indiana University or Syracuse University participated for a class requirement or $10 per hour.

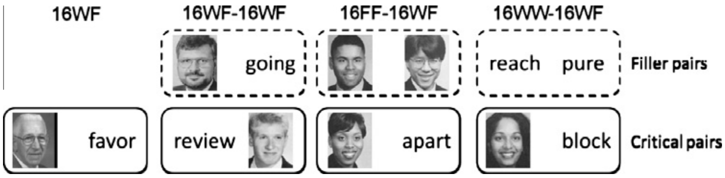

Fig. 5. The design employed in Experiment 5. The filler pairs were either word–word, face–face, or word–face pairs added to the beginning of the study list. The critical pairs were the 16 word–face pairs that appeared at the end of the list in each condition. The critical pairs were the only pairs that appeared on a subsequent cued recall test.

Materials

The materials used were the same as those described in Experiment 1.

Design and procedure

Participants were randomly assigned to one of four between-subject study conditions: 16WF pairs, 16WF–16WF pairs, 16FF–16WF pairs, or 16WW–16WF pairs. The experimental design is shown in Fig. 5. Study was incidental and each study list contained 16 critical WF pairs that occurred in the final 16 trials of the study list. These are the pairs that will be tested (unbeknownst to the participants during study). The length of the study list was manipulated by adding pairs prior to the critical pairs. Participants in the 16WF–16WF condition studied 16 additional WF pairs. Participants in the 16FF–16WF condition studied 16 FF pairs and participants in the 16WW–16WF condition studied 16 WW pairs. This design ensured that study-test lag was constant across conditions and that participants would treat the items as a single series. During study, participants saw each pair for 3 s and were asked to rate the degree of association between the items. Participants were tested on the 16 critical WF pairs. During the test participants were provided with a face and asked to recall the word it was presented with during study with responses coded as described in Experiment 1.

Results and discussion

A one-way ANOVA performed on the correct recall data revealed a significant effect of condition on correct recall ($F(3,202) = 4.47$, $p = .005$, $\eta_p^2 = .063$). Next we performed a Dunnett's two-sided *t*-test to compare the conditions where pairs were added to the list (i.e., 16WW–16WF, 16FF–16WF, 16WF–16WF) to our baseline 16WF condition in a single analysis. Dunnett's test allows for the comparison of experimental conditions to a baseline while controlling for family-wise error rate (Dunnett, 1955, 1965). The results of the test revealed that only the 16WF–16WF ($p = .046$, Cohen's $d = .167$) group differed significantly from the 16WF group. Thus adding WW (Cohen's $d = .054$) and FF (Cohen's $d = .095$) pairs to the study list did not harm performance relative to the 16WF pair condition, but adding an additional 16WF pairs to the list did. No effect of condition was found for intrusions ($F < 1$) or no response trials ($F(3,199) = 2.44$, $p = .066$, $\eta_p^2 = .035$). The means and standard error for all response categories and all conditions are presented in Table 4.

Table 4
The probability of each response type by study condition for Experiment 5.

	P (Correct)		*P* (Intrusion)		*P* (No recall)	
	M	*SE*	*M*	*SE*	*M*	*SE*
16WF	.225	.024	.170	.025	.605	.027
16WF–16WF	.156	.017	.221	.020	.623	.022
16FF–16WF	.264	.021	.214	.032	.523	.033
16WW–16WF	.248	.029	.196	.032	.556	.033

We observed a list length effect specific to pair-type. Adding pairs of the same type as the tested pairs (i.e., WF pairs) but not pairs of different types harmed performance. These data, along with findings of Experiments 1 and 2, reinforces the need for distinct associative features. Following Criss and Shiffrin (2004, 2005) we suggest that participants use class attributes to restrict memory search to the relevant subset of memory and do so flexibly for associative (but not item) tasks.

General discussion

The purpose of the current paper was to better understand the representation of associative information in memory. We adopted the paradigm Criss and Shiffrin (2004, 2005) used to study associative information in associative recognition and applied it to cued recall. In Experiments 1 and 2 we found that items studied in List 1 influenced subsequent performance for a List 2 memory test only if those List 1 items were composed of the same pair-type as List 2. Critically, when the same design was employed but the test was changed to single item recognition instead of cued recall, List 1 exerted the same influence on memory for List 2 items regardless of the class of pairs that appeared in the first list. Importantly the encoding task, instructions, and expectations were identical for the single item and cued recall experiments. Therefore the process responsible for the different pattern in cued recall and single item recognition must be at retrieval. In the final experiment, we demonstrated a within pair-type list length effect but no across pair-type list length effect in cued recall. There are two critical findings that need explanation. First, the pair-type effects found in tests for associative memory (e.g., cued recall here and associative recognition in Criss and Shiffrin (2004, 2005)) but not single item memory. Second, the increase in correct and incorrect responses to cues that appeared twice during study in Experiment 1.

Associative information

The pair-type effects are consistent with associative recognition findings and provide compelling evidence for distinct associative representations, but are problematic for current models of memory. Let's consider a simple model for cued recall based on the REM framework representing associative information as a simple co-occurrence of item information (e.g., Diller, Nobel, & Shiffrin, 2001). During study, the two items from each pair are encoded together in a single memory trace. At test, the presented cue is compared to episodic memory traces and a match to each memory trace is computed just as in single item recognition. A single memory trace is sampled in proportion to how well it matches the cue. If that match exceeds a threshold, the participant accepts that trace as a candidate and attempts to retrieve and output the target contained in the sampled trace. If successful, a response is emitted. If not, the process repeats until a number of failures is reached and search is then terminated (e.g., Harbison, Dougherty, Davelaar, & Fayyad, 2009; Mensink & Raaijmakers, 1988; Raaijmakers & Shiffrin, 1981).

If associative information were represented by the simple co-occurrence of item information then adding pairs of any type or repeating items in any type would cause interference (e.g., increasing the noise in the matching and/or sampling process). The design of Experiments 1 and 2 ensures that the total number and type of single items was identical, only the composition of the pairs differed. Only when the composition of List 1 matched List 2 did we find interference. We could potentially account for the increase in correct and incorrect responses for the rearranged pairs of Experiment 1 by assuming that participants increase the number of failures required to terminate search or decrease their threshold for accepting a sampled trace. However either of these assumptions would also predict the same pattern of data in Experiment 2. To account for the different patterns of data, we must assume that different information contributes to cued recall depending on the pair-type. The design of Experiment 5 allowed us to vary the number of items that were similar to the cue (e.g., adding 16 FF pairs) or similar to the target (e.g., adding 16 WW pairs). In neither case did accuracy suffer. Only when adding pairs of the same type as those being tested did we find a list length effect. Together, these data are clearly incompatible with co-occurrence models. Instead, participants strategically used associative information, we suggest class attributes, to limit their search of memory to relevant information. We suggest that participants include class attributes in the memory probe whenever attempting to retrieve memories for an association. This is an effective strategy because using class attributes eliminates interference from some of the studied items (e.g., those with different class attributes).

Why participants did not use pair-specific class attributes when attempting to retrieve memories for items is a bit of a puzzle. The study designs were identical so using class attributes would also eliminate interference in single item recognition as was the case for cued recall. However, participants did not use this information to their advantage. Perhaps participants simply do not employ class attributes because nothing about the cue or the task in single item recognition requires knowledge about the pair. This would avoid wasting already limited cognitive resources to invoke information that is not needed (e.g., Gillund & Shiffrin, 1984; Murdock, 1997; Raaijmakers & Shiffrin, 1981). This highlights the fact that memory is incredibly flexible and the cues used to retrieve memories are adapted to best meet the demands on the memory system (e.g., Humphreys, Bain, & Pike, 1989; Humphreys et al., 2003; Malmberg & Xu, 2007).

Lastly, consider emergent features and associative links. The current data tell us little about such features. Emergent features (and associative links) are not part of the cue in cued recall and thus cannot be used to search memory. In fact, only when successfully generating a response are such features available. However, we find the many experiments cited earlier, and the Clark and Gronlund (1996) review, provide compelling evidence in favor of the existence of such features. However, the current data are problematic even for models that represent associative information using emergent features (e.g., Murdock, 1982, 1997) because the memory vector is composite and contains information from all pair-types. Consequently, these models predict that all studied items regardless of pair-type cause interference (see Criss and Shiffrin (2004) for further discussion). Thus, even in an emergent features model, class attribute information is necessary to restrict interference to a subset of the emergent features, namely those from the same type of pair in an associative memory task. Underwood (1969) conceived of class attributes as a set of general features, like item features, that contribute to memory. Similarly, we think that class attributes are amenable to all types of models regardless of how the model represents pair-specific associative information (e.g., emergent features or co-occurrence).

Increase in correct and incorrect responses for rearranged cues

In Experiment 1 we found an increase in both correct and incorrect responses for cues that were encoded twice (in two different pairs on two different lists), compared to cues encoded just once. Though this is not the focus of the paper, we found this pattern of data intriguing. As such, the following discussion is not intended to explain the pattern of data but instead to speculate.

First we note that the current finding appears contrary to proactive interference in paired associate designs. Proactive interference is typically observed when a cue appears with two different responses across lists compared to a cue that appeared with just one response (e.g., Postman & Gray, 1977; Postman, Stark, & Burns, 1974; see Anderson and Neely (1996) for a review). However, the current methodology is much different. The current design uses WF rather than WW pairs, tests both conditions (cues with 1 vs. 2 responses) in a single list (rather than the standard between-subject design for paired associates), and the pairs are studied just once under incidental conditions rather than practiced until reaching a criterion level of accuracy. Any combination of these factors, especially the use of single trial learning rather than training to a criterion, may be the reason we do not observe the standard proactive interference effect. We consider three possible explanations for our data: 1) longer search time for rearranged cues than List 2 only cues 2) lower thresholds for accepting a sampled trace for rearranged vs. List 2 only cues or 3) enhanced encoding of rearranged pairs.

First consider search time. Mensink and Raaijmakers (1988, see Fig. 7) simulated a paired associate task with a control condition (similar to List 2 only in our experiment) and an interference condition (similar to rearranged in our experiment). They varied the number of attempts at retrieval. With few attempts recall was better for the control condition, but with many attempts recall was better for the interference condition. Diller et al. (2001) assumed that the number of retrieval attempts (e.g., the number of failed attempts one allows before stopping the search) is sensitive to cue familiarity: the more familiar the cue, the longer participants are willing to search. Recall that in the current experiments, rearranged pairs are composed of items studied in different pairs on each study list, thus the individual items are repeated, but the pair itself was studied just once. Thus a rearranged cue has a higher level of overall familiarity in the experimental context than a cue from the List 2 only pairs. Together, the Mensink and Raaijmakers and Diller et al. simulations suggest that longer search increases response rates and that cue familiarity drives search time. Applying this logic to our paradigm, we speculate that rearranged (e.g., familiar) cues lead to longer search which increases the rate of correct responses (as in Mensink and Raaijmakers) and the rate of incorrect responses.

Another possibility is that participants lower their threshold for accepting a sampled trace in response to familiar cue, resulting in more attempts to recover and generate a target. For example, participants may feel that they ought to remember the target word, given the high familiarity evoked by the cue and output a response that they would otherwise have withheld. Additional support for this idea comes from research examining what people know about the accuracy of their future memory (i.e., feeling of knowing judgments). For cued recall these judgments are driven, in part, by the familiarity of the cue (Benjamin, 2005; Koriat & Levy-Sadot, 2001; Metcalfe, Schwartz, & Joaquim, 1993; Reder, 1987; Schwartz & Metcalfe, 1992). Likewise, in the current experiments, participants may feel confident that they remember the target based on the familiarity of the cue and are therefore more likely to give a response. One strategy to test this hypothesis is to force participants to respond on each trial (e.g., Koriat & Goldsmith, 1994, 1996). In a forced report paradigm, differential thresholds should be ruled out as participants are required to generate a response to every cue. Both of these cases, changing the threshold for accepting a sampled trace or changing the search termination rule, are reminiscent of response bias. In both cases, the accuracy of the encoded memory traces do not differ for a rearranged vs. List 2 only conditions, instead the basis for responding has changed. We now consider the converse, perhaps repetition of individual items improves encoding of the pair.

Perhaps the second presentation of an item in a different pair than the initial presentation results in better encoding for the second pair. It is conceivable that the study task for the second study list (i.e., generate a sentence about the items) could have facilitated the use of the first list target as a mediator. Similarly, Howard, Jing, Rao, Provyn, and Datey (2009) suggest that such associations could be driven by presentation of the items during a similar temporal context. After initially learning A–B, when the participant is shown A–D part of the temporal context associated with A (i.e., B) is reinstated resulting in a transitive association between B and D. In both instances, a participant may have more information than that provided by the cue alone. Clark and Burchett (1994) and Kahana and Caplan (2002) both demonstrated that probing with two parts of a studied triplet (i.e., compound cues) improved correct recall, and the same could be the case here. For instance, a participant provided with A could recall B. If they recalled that B was part of the first list, they could then use A–B to search memory for D. If they failed to recall B was part of the first list, then they may report B which would be coded as an intrusion. Finally, we note that half of the stimuli were novel faces, never before seen by our participants. Perhaps it is more difficult to form an association between unfamiliar stimuli (e.g., Xu & Malmberg, 2007) and this difficulty is partially alleviated with repetition of the unfamiliar stimuli. Thus, items repeated on List 2 and pairs containing those repeated items are better encoded simply because the novel items are more familiar following presentation on List 1. Whether the increase in correct and incorrect responses for rearranged pairs is due to increase leniency at retrieval, increased search attempts, or enhanced encoding (or some other mechanism) is unknown. This question deserves attention in future research.

Summary

The results of the current paper indicate that the storage of associative information beyond the mere co-occurrence of item information is necessary in developing a model that can successfully account for associative recognition, cued recall, and single item recognition data. We suggest that these features may take the form of emergent features and class attributes. Emergent features describe the unique combination of two items (e.g., a gestalt). Class attributes, as proposed by Underwood (1969), define the class of information. The specific set of associative features used depends on the information provided by the cue and the task. In associative recognition, emergent features and class attributes are part of the cue and are used to probe memory (e.g., see Criss and Shiffrin (2004, 2005)). In cued recall, class attributes are the only associative features used to probe memory. The cued recall data also showed that participants are more likely to respond to familiar cues, perhaps because they spend more time searching their memory in response to familiar cues or perhaps because repetition of the individual items increases the accuracy of the encoded pair. These experiments reinforce the proposal that people flexibly and strategically use cues to search memory based on task demands.

Acknowledgments

This research was supported by National Science Foundation Grant# 0951612 to AHC. The authors would like to thank Marc Howard for his input during the drafting of this manuscript and Stephanie Cuevas for skillfully collecting and coding the data. The authors would also like to thank Jeroen Raaijmakers, Jim Neely, and an anonymous reviewer for their insightful comments on the manuscript.

References

Anderson, M. C., & Neely, J. H. (1996). Interference and inhibition in memory retrieval. In E. Bjork, R. A. Bjork, E. Bjork, & R. A. Bjork (Eds.), *Memory* (pp. 237–313). San Diego, CA, US: Academic Press.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., et al. (2007). The English lexicon project. *Behavior Research Methods, 39*, 445–459.

Benjamin, A. S. (2005). Response speeding mediates the contributions of cue familiarity and target retrievability to metamnemonic judgments. *Psychonomic Bulletin & Review, 12*, 874–879.

Clark, S. E., & Burchett, R. R. (1994). Word frequency and list composition effects in associative recognition and recall. *Memory & Cognition, 22*, 55–62.

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review, 3*, 37–60.

Criss, A. H. (2005). The representation of single items and associations in episodic memory. (Doctoral dissertation, Indiana University, 2004). *Dissertation Abstracts International B, 65*, 6882.

Criss, A. H., & Shiffrin, R. M. (2004). Pairs do not suffer interference from other types of pairs or single items in associative recognition. *Memory & Cognition, 32*, 1284–1297.

Criss, A. H., & Shiffrin, R. M. (2005). List discrimination in associative recognition and implications for representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 1199–1212.

Criss, A. H., Aue, W. R., & Smith, L. (2011). The effects of word frequency and context variability in cued recall. *Journal of Memory and Language, 64*, 119–132.

Diana, R. A., Reder, L. M., Arndt, J., & Park, H. (2006). Models of recognition: A review of arguments in favor of a dual-process account. *Psychonomic Bulletin & Review, 13*, 1–21.

Diller, D. E., Nobel, P. A., & Shiffrin, R. M. (2001). An ARC–REM model for accuracy and response time in recognition and recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 414–435.

Dunnett, C. W. (1955). A multiple comparison procedure for comparing several treatment groups with a control. *Journal of the American Statistical Association, 50*, 1096–1121.

Dunnett, C. W. (1965). New tables for multiple comparisons with a control. *Biometrics, 20*, 482–491.

Eich, J. M. (1982). A composite holographic associative recall model. *Psychological Review, 89*, 627–661.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91*, 1–67.

Gronlund, S. D., & Ratcliff, R. (1989). Time course of item and associative information: Implications for global memory models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 846–858.

Harbison, J., Dougherty, M. R., Davelaar, E. J., & Fayyad, B. (2009). On the lawfulness of the decision to terminate memory search. *Cognition, 111*, 416–421.

Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments & Computers, 16*, 96–101.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review, 95*, 528–551.

Hockley, W. E., & Consoli, A. (1999). Familiarity and recollection in item and associative recognition. *Memory & Cognition, 27*, 657–664.

Hockley, W. E., & Cristi, C. (1996a). Tests of encoding tradeoffs between item and associative information. *Memory & Cognition, 24*, 202–216.

Hockley, W. E., & Cristi, C. (1996b). Tests of the separate retrieval of item and associative information using a frequency-judgement task. *Memory & Cognition, 24*, 796–811.

Howard, M. W., Jing, B., Rao, V. A., Provyn, J. P., & Datey, A. V. (2009). Bridging the gap: Transitive associations between items presented in similar temporal contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 391–407.

Humphreys, M. S., Bain, J. D., & Pike, R. (1989). Different ways to cue a coherent memory system: A theory for episodic, semantic and procedural tasks. *Psychological Review, 96*, 208–233.

Humphreys, M. S., Dennis, S., Maguire, A. M., Reynolds, K., Bolland, S. W., & Hughes, J. D. (2003). What you get out of memory depends on the question you ask. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 797–812.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*, 513–541.

Kahana, M. J., & Caplan, J. B. (2002). Associative asymmetry in probed recall of serial lists. *Memory & Cognition, 30*, 841–849.

Kelley, R., & Wixted, J. T. (2001). On the nature of associative information in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 701–722.

Koriat, A., & Goldsmith, M. (1994). Memory in naturalistic and laboratory contexts: Distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General, 123*, 297–315.

Koriat, A., & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103*, 490–517.

Koriat, A., & Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 34–53.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present day American English*. Providence, RI: Brown University Press.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments & Computers, 28*, 203–208.

Malmberg, K. J., & Xu, J. (2007). On the flexibility and the fallibility of associative memory. *Memory & Cognition, 35*, 545–556.

McClelland, J. L., & Chappell, M. (1998). Familiarity breeds differentiation: A subjective-likelihood approach to the effects of experience in recognition memory. *Psychological Review, 105*, 724–760.

Mensink, G., & Raaijmakers, J. G. (1988). A model for interference and forgetting. *Psychological Review, 95*, 434–455.

Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 851–864.

Murdock, B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review, 89*, 609–626.

Murdock, B. (1997). Context and mediators in a theory of distributed associative memory (TODAM2). *Psychological Review, 104*, 839–862.

Murnane, K., Phelps, M. P., & Malmberg, K. (1999). Context-dependent recognition memory: The ICE theory. *Journal of Experimental Psychology: General, 128*, 403–415.

Murnane, K., & Shiffrin, R. M. (1991). Word repetitions in sentence recognition. *Memory & Cognition, 19*, 119–130.

Nobel, P. A., & Shiffrin, R. M. (2001). Retrieval processes in recognition and cued recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 384–413.

Overman, A. A., & Becker, J. T. (2009). The associative deficit in older adult memory: Recognition of pairs is not improved by repetition. *Psychology and Aging, 24*, 501–506.

Postman, L., & Gray, W. (1977). Maintenance of prior associations and proactive inhibition. *Journal of Experimental Psychology: Human Learning and Memory, 3*, 255–263.

Postman, L., Stark, K., & Burns, S. (1974). Sources of proactive inhibition on unpaced tests of retention. *The American Journal of Psychology, 87*, 33–56.

Raaijmakers, J. G., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review, 88*, 93–134.

Reder, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology, 19*, 90–138.

Reder, L. M., Nhouyvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember-know judgments in a continuous recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 294–320.

Schwartz, B. L., & Metcalfe, J. (1992). Cue familiarity but not target retrievability enhances feeling-of-knowing judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 1074–1083.

Shiffrin, R. M., Murnane, K., Gronlund, S., & Roth, M. (1989). On units of storage and retrieval. In C. Izawa & C. Izawa (Eds.), *Current issues in cognitive processes: The Tulane Floweree symposium on cognition* (pp. 25–68). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM – Retrieving effectively from memory. *Psychonomic Bulletin & Review, 4*, 145–166.

Underwood, B. J. (1969). Attributes of memory. *Psychological Review, 76*, 559–573.

Underwood, B. J., & Schulz, R. W. (1960). *Meaningfulness and verbal learning*. Oxford, England: J.B. Lippincott.

Xu, J., & Malmberg, K. J. (2007). Modeling the effects of verbal- and nonverbal-pair strength on associative recognition. *Memory & Cognition, 35*, 526–544.