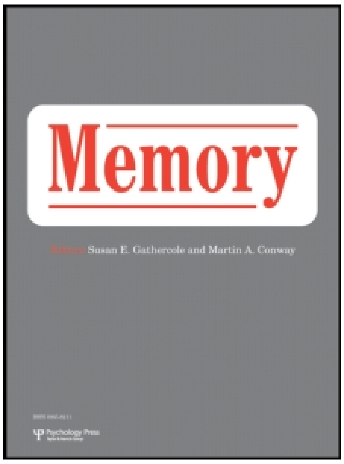

Memory
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713683358

Contributions of familiarity and recollection rejection to recognition: Evidence from the time course of false recognition for semantic and conjunction lures

Laura E. Matzen[a]; Eric G. Taylor[b]; Aaron S. Benjamin[b]
[a] Sandia National Laboratories, Albuquerque, NM, USA [b] University of Illinois, Urbana-Champaign, IL, USA

Online publication date: 13 January 2011

To cite this Article Matzen, Laura E. , Taylor, Eric G. and Benjamin, Aaron S.(2011) 'Contributions of familiarity and recollection rejection to recognition: Evidence from the time course of false recognition for semantic and conjunction lures', Memory, 19: 1, 1 — 16
To link to this Article: DOI: 10.1080/09658211.2010.530271
URL: http://dx.doi.org/10.1080/09658211.2010.530271

MEMORY, 2011, 19 (1), 1–16

Psychology Press
Taylor & Francis Group

Contributions of familiarity and recollection rejection to recognition: Evidence from the time course of false recognition for semantic and conjunction lures

Laura E. Matzen[1], **Eric G. Taylor**[2], **and Aaron S. Benjamin**[2]

[1]Sandia National Laboratories, Albuquerque, NM, USA
[2]University of Illinois, Urbana-Champaign, IL, USA

It has been suggested that both familiarity and recollection contribute to the recognition decision process. In this paper we leverage the form of false alarm rate functions—in which false alarm rates describe an inverted U-shaped function as the time between study and test increases—to assess how these processes support retention of semantic and surface form information from previously studied words. We directly compare the maxima of these functions for lures that are semantically related and lures that are related by surface form to previously studied material. This analysis reveals a more rapid loss of access to surface form than to semantic information. To separate the contributions of item familiarity and reminding-induced recollection rejection to this effect, we use a simple multinomial process model; this analysis reveals that this loss of access reflects both a more rapid loss of familiarity and lower rates of recollection for surface form information.

Keywords: Recognition; Familiarity; Reminding; Recollection rejection; Conjunction lures; Semantic lures; Multinomial process modelling.

Recent memory research incorporates a prominent role for understanding the ways in which misremembering, as well as remembering, can inform memory theory (Gallo, 2006; Roediger, 1996). Within recognition memory, evaluating the relationships between studied stimuli and test stimuli and the conditions that promote false remembering reveals a great deal about the encoding strategies (Matzen & Benjamin, 2009), decision processes (Benjamin, 2001; Miller & Wolford, 1999), and representations that support recognition (cf. Brainerd & Reyna, 2001; Brainerd, Reyna, Wright, & Mojardin, 2003).

This paper adds to the literature on memory errors by using them to investigate the time courses of two component processes that are thought to contribute to recognition—familiarity and recollection rejection. Specifically, we compare these two processes for lures that are semantically related to studied words as opposed to lures that share surface features (syllables with the same orthography) but not semantic features with studied words. Previous research suggests that differences in how semantic and surface form information are processed at study (Matzen & Benjamin, 2009) and at test (Odegard, Lampinen, & Toglia, 2005) play a crucial role in determining how susceptible people will be to different types of memory errors. Comparing the time courses of familiarity and recollection rejection for lures that

Address correspondence to: Laura E. Matzen, Cognitive Systems Research and Applications, Sandia National Laboratories, P.O. Box 5800, Albuquerque, NM 87185-1011, USA. E-mail: lematze@sandia.gov

This work was funded in part by grant R01 AG026263 to Aaron Benjamin from The National Institutes of Health. It was also supported by the Sandia National Laboratories Excellence in Engineering Fellowship provided to Laura Matzen and is a part of her doctoral dissertation.

http://www.psypress.com/memory DOI:10.1080/09658211.2010.530271

are related to studied items by shared semantics or shared surface features allows us to draw novel conclusions about the time course of item information decay and the processes underlying that decay. Through experimental evidence and modelling we lay out a theory about the relationship between familiarity and recollection that accounts for the joint set of false alarm functions observed for two different types of lures.

Conjunction lures and semantic lures have been used in numerous studies of verbal memory. Conjunction lures are test items that have high orthographic overlap but generally little semantic overlap with studied items. They are typically created by combining morphemes from two studied compound words, such as ''blackmail'' and ''jailbird'', to form a test word such as ''blackbird''. Participants often endorse these lures at high rates (Jones & Atchley, 2006; Jones & Jacoby, 2001; Marsh, Hicks, & Davis, 2002; Reinitz, Lammers, & Cochran, 1992; Underwood, Kapelak, & Malmi, 1976; Underwood & Zimmerman, 1973), particularly when the study context tacitly encourages some retention of surface structure (Benjamin, 2008; Matzen & Benjamin, 2009).

Semantic lures are test items that have high semantic overlap and little or no orthographic overlap with studied words. They are typically synonyms or close semantic associates, such as ''bunny'' and ''rabbit''. Semantic errors, or cases in which participants endorse semantic lures or incorrectly recall semantic associates, have been seen frequently in experiments in which the studied items are sentences or longer texts (Bock & Brewer, 1974; Bransford & Franks, 1971; Brewer, 1977; Johnson, Bransford, & Solomon, 1973; Matzen & Benjamin, 2009), in lists containing multiple words from the same semantic category (Heathcote, 2003, Shiffrin, Huber & Marinelli, 1995) and in experiments using the DRM paradigm, which involves presenting a large number of semantic associates during study (Deese, 1959; McDermott & Watson, 2001; Roediger & McDermott, 1995; Roediger, McDermott & Robinson, 1998). In the present experiment we used both conjunction and semantic lures and varied the lag between the presentation of studied items and their associated lures. In doing so we were able to track different factors that contribute to memory errors and how they change as the time between study and test increases.

THEORETICAL MECHANISMS UNDERLYING MEMORY ERRORS

Two theories that are commonly used to account for semantic and conjunction errors are dual-process theory (cf. Jones & Jacoby, 2001; Yonelinas, 2002) and fuzzy-trace theory (cf. Brainerd & Reyna, 2001; Brainerd et al., 2003). The theories account for memory errors in fundamentally similar ways. The dual-process theory posits that memory errors occur when similarities between a test item and one or more studied items create a misleading amount of familiarity for the test item. If the participant is also unable to recall the related studied items and does not realise that they are in fact different from the lure, this level of familiarity leads participant to endorse the lure (Jacoby, 1996; Jones & Atchley, 2002, 2006; Jones & Jacoby, 2001; Lampinen, Odegard, & Neuschatz, 2004). In dual-process theory the processes leading both to high levels of familiarity and to recollection can occur with either the semantic or surface form information that the participant encoded from a studied item (Jacoby, 1996). Fuzzy-trace theory is generally similar to the dual-process theory, but it additionally specifies that only gist traces (memory for the general meaning of an item) produce familiarity while verbatim traces (memory for the surface form of an item) support recollection. In order to reject a gist-consistent lure people must be able to retrieve a verbatim trace, which stores surface form information (Brainerd & Reyna, 1998, 2001, 2007; Brainerd et al., 2003).

Both the dual-process and the fuzzy-trace theory include a process by which participants are able to reject lures by recalling the parent items from which the lures were derived. This process, called recollection rejection, allows the participant to discount the sense of familiarity elicited by a lure. Evidence of recollection rejection has been found for both conjunction lures (Jones, 2005; Jones & Atchley, 2006; Jones & Jacoby, 2001; Lampinen et al., 2004) and for semantic lures (Brainerd et al., 2003). Throughout this paper we will use the term *familiarity* to refer to the degree to which a stimulus elicits generic evidence for having been previously studied, and *recollection* for the process by which the specific prior encounter that promoted that level of familiarity is accessed. We use the term *recollection rejection* to refer to cases in which the process of recollection for a specific studied item provides participants with information that they can use to reject a related lure.

NON-MONOTONIC FALSE ALARM RATES

Experiments that manipulate the interval between study and test have revealed a non-monotonic false alarm function for both semantic (Brainerd et al., 2003; MacLeod & Nelson, 1976; see also Ratcliff & Hockley, 1980) and conjunction lures (Jones & Atchley, 2002, 2006). In a non-monotonic false alarm function, false alarm rates are low when the lag between study and test is short (or when the study time is very short in experiments that manipulate presentation duration rather than lag). The false alarm rates increase over intermediate lags and decrease again at long lags, creating an inverse U shape. A non-monotonic function reveals the interactivity of two opposing processes with different time courses. For false alarms it is thought that the two opposing processes are familiarity and recollection. When the time between study and test is short, good memory for the studied stimulus leads the lures to be highly familiar, but it also affords a high rate of recollection rejection, allowing participants to account for this familiarity and discount the lure appropriately. Consequently the participants endorse few lures at short lags. As the time between study and test increases, the rate of recollection and the familiarity of the lures both decrease, but because the representations that promote familiarity have a longer half-life than those that promote recollection (Benjamin & Bjork, 2000), the false alarm rate increases. The participants are no longer as successful at countering the familiarity by recollecting the originally studied items. When there is a long lag between study and test the false alarm rate decreases once again. Although there is little recollection rejection occurring at these long lags, the lures themselves are also less likely to be familiar as more and more of the information about the studied items is forgotten.

While this general pattern has been observed for both conjunction and semantic lures, the relationship between the lure and its parents seems to play an important role in the process of recollection rejection. When conjunction lures share semantic overlap as well as orthographic overlap with their parent items (such as the lure "handgun" drawn from the parents "handball" and "shotgun"), participants report using recollection rejection at much higher rates than they do when the conjunction lures are related to their parents only by surface form (Odegard et al., 2005). Similarly, Odegard and Lampinen (2005) found higher rates of recollection rejection for lures that were antonyms of the studied words than for words that rhymed with the studied words. These results indicate that lures that overlap with the studied items' surface form but do not overlap in meaning may not be adequate cues for retrieval of the studied items. In other words, semantic similarity can facilitate recollection rejection while surface form similarity is much less likely to support this process. Although the similar meanings of the parents and targets make the lures more appealing, the semantically similar lures can also act as cues to help the participants recall the original parent items, thus enabling them to reject the lure. In that sense, the test items can be thought of as reminders of the study events that promoted their familiarity (cf. Bellezza, Winkler, & Andrasik, 1975; Bray & Robbins, 1976; Hintzman, 1976). When the test item is an efficient reminder, recollection rejection is more likely than when it is not.

THE TIME COURSE OF FORGETTING OF SURFACE FORM AND SEMANTIC INFORMATION

There are reasons to suspect that the retention of surface form and semantic information may follow different time courses. Previous research has shown that when participants encode a word, they are generally likely to retain more information about its meaning than about its exact form (Bock & Brewer, 1974; Brewer, 1975; Matzen & Benjamin, 2009; Potter & Lombardi, 1990; Sachs, 1967). Models of reading typically presume that lower-level information is discarded once meaning has been extracted from a message. This could be due to qualitative differences in how surface form information and semantic information are encoded or, as we assume here, readers may simply encode less surface form information than semantic information. As the time between study and test increases, information about the studied word is forgotten. If there was less information about the surface form of the study word encoded initially, this information will be rendered unrecoverable more quickly than the more redundantly encoded semantic information (Benjamin, 2008). Consequently there should be a more rapid loss of familiarity for surface form information than for semantic information.

In addition, the alphabet of surface forms (literally, the alphabet) is much more restricted than the alphabet of meaning. Therefore it would not be surprising for interference to be greater at lower levels, such as between letters and even morphemes, than between words themselves. The finding that nonwords are judged to be more similar than words to a previously studied mix of nonwords and words supports this notion (Greene, 2004). Differences in interference levels for surface form and semantic information as well as differences in the amount of each type of information encoded at study are both factors that could lead to differences in the rate at which access to semantic and surface form information is lost.

Independent of the decay functions for the surface and semantic features of studied words, there are also likely to be differences in the rates of recollection rejection for semantic and conjunction lures. As discussed above, lures that share semantic overlap with their parent items are more likely to cue recollection than lures that have only surface form overlap (Odegard & Lampinen, 2005; Odegard et al., 2005; Thios & D'Agostino, 1976). This could be due to greater interference for surface form information than for semantic information or greater distinctiveness for semantic information (Gallo, Meadow, Johnson & Foster, 2008; Jacoby & Craik, 1979; Lloyd, 2007; Moscovitch & Craik, 1976). In either case, if there are baseline differences in the abilities of semantic and surface form information to cue recollection, those differences will be exacerbated as the time between study and test increases. For example, when there are many intervening words between the studied item and the test item there is likely to be a great deal of orthographic overlap across the intervening words, but those words may share few if any semantic features with the previously studied word. Thus, as the time between study and test increases, the effectiveness of surface form information as a cue for recollection is likely to decrease much faster than the effectiveness of semantic information.

In summary, while there are good a priori reasons to predict a more rapid loss of access to surface form than to semantic information, the factors that might drive that loss of access are unclear. It may simply be that surface form information is forgotten more quickly than semantic information, either by virtue of greater interference or reduced ongoing maintenance processes. Alternatively, test items that are semantically related to studied items may be superior cues for the recovery of the studied items than are test items that share only surface overlap with studied items (e.g., Thios & D'Agostino, 1976). In that case, memory for the original surface form and semantic information may be equivalent, but semantic information is more likely to be effectively cued by later events than is surface form information.

In this paper we use the degree to which lures induce false remembering as a means to evaluate these questions about what it means for access to be lost. The false alarm rate function reflects the balance between the familiarity of a lure and the participant's ability to reject the lure by recollecting information about the originally studied items. The similarity between a previously studied item and a lure should promote endorsement of that lure, but at the same time, if the lure cues successful recollection of the related study item, the lure will be rejected. As the time between study and test increases, the balance between these two opposing processes shifts, leading to changes in the false alarm rate. However, as discussed above, the loss of familiarity and the ability to cue recollection rejection are likely to differ for semantic and surface form information. Here we use a combination of methodologies to tease apart the effects of differential forgetting and reminding for these two types of information.

To date, no experiments have directly compared the non-monotonic false alarm functions for conjunction and semantic lures. By making this comparison we tested two related hypotheses: (1) that the familiarity yielded by surface form representations decreases faster over time than the familiarity of semantic information, and (2) that semantic information serves as a better reminder for countering familiarity through recollection rejection. Both of these factors could affect the location of the peak of the non-monotonicity in the false alarm functions. Consequently, if familiarity and recollection rejection contribute differentially to the recognition decision process for semantic and conjunction lures, then the peak of the non-monotonicity will occur at different points in time for these two types of lures. Our prediction in this experiment was that a combination of differences in familiarity and differences in recollection rejection would cause the false alarm function to peak at later lags for semantic lures than for conjunction lures. We chose to compare the two false alarm functions based on the location of the peak of the non-monotonicity because our primary focus is on the trade-off between familiarity and recognition in the participants' decision-making processes. In this type of comparison the relative locations of the peaks for the two functions is the crucial information, and any differences in the overall false alarm rates for the two types of lures are irrelevant to the analysis. This eliminated the need to equate overall performance across the two conditions.

In order to investigate the differential contributions of familiarity and recollection rejection to the false alarm functions in the present study, we used a continuous recognition task with a set of four possible response choices that were similar to remember–know judgements. In a typical experiment using remember–know judgements, participants in a recognition memory task are asked to make a judgement when they endorse an item as being an old, previously studied item. They are asked to indicate whether they endorsed the item because they specifically remember seeing it at study (the ''remember'' judgement) or whether they endorsed the item because they have a sense that it is familiar, even if they don't specifically remember seeing it at study (the ''know'' judgement; Tulving, 1985). These judgements can be interpreted as reflecting subjectively different states of knowledge or as different sides of a response criterion reflecting different levels of confidence (cf. Benjamin, 2005; Dunn, 2004).

In this experiment we made the assumption that the differences between responses reflect the participants' placement of response criteria rather than reflecting different states of knowledge or different retrieval processes. We are interested in the contributions of recollection and familiarity to both false alarms to lures and correct rejections of lures. We take ''remember'' to be a response indicating that a participant recollects (correctly or incorrectly) enough details about an item to make a high-confidence judgement that the item was in the study list. We take the ''know'' response to indicate that the participant believes that an item was in the study list, but has low confidence in that decision, perhaps because he or she cannot recollect details about the studied item. We were interested in the contributions of recollection and familiarity to both ''yes'' and ''no'' responses, so we used a procedure in which participants were asked to make a judgement similar to the remember–know judgement for all of the test items, regardless of whether they thought the items were new or old. If the participants saw a test item that they thought was an old, previously studied item, they were asked to select one of two response choices, either ''remember'' or ''familiar''. They were instructed to select ''remember'' if they specifically remembered studying the word earlier in the experiment, or ''familiar'' if they did not specifically remember studying it but thought that it seemed familiar and was likely to have been a studied item. The ''familiar'' response is equivalent to the more commonly used ''know'' response, but a different term was chosen in order to make the meaning of this response choice clearer to the participants. If the participants saw a test word that they thought was a new, unstudied word, they were asked to respond by selecting either ''different'' or ''unfamiliar''. These responses are similar to those used in Jones and Atchley (2006) in their investigation of recollection rejection for conjunction lures. The participants in the present study were instructed that some of the test words would be words that were very similar to one or more studied words, but that were actually different words. They were asked to use the ''different'' response in cases where they recognised that a test word was a new word because they were able to remember a similar but slightly different word from the study phase. The participants' ''different'' responses were used as a measure of recollection rejection (much like the ''recollect'' response in Jones & Atchley, 2006). The ''unfamiliar'' response was more of a generic ''no'' response, and participants were asked to use this response when they encountered test words that they thought were new and they did not specifically remember seeing a similar word at test. Note that the instructions to participants intentionally encouraged them to use a recall-to-reject strategy, which was crucial for our comparison of the differential contributions of familiarity and recollection to the false alarm functions for semantic and conjunction lures.

The participants' responses in each of these four categories were used to determine false alarm rates and the rates of recollection rejection at each lag. This information can be used directly to test our predictions about different rates of recollection rejection for the two types of lures. In addition we used the response patterns at each lag in order to fit a simple multinomial model that represents the participants' use of recollection or familiarity in making decisions about each lure type.

METHOD

Participants

A total of 58 University of Illinois undergraduates (35 female) participated in the experiment for credit in an introductory psychology course. The mean age of the participants was 19 (range 18–22).

Design

The study employed a 9 (lag) × 4 (test item type: old, conjunction lure, semantic lure, unrelated new item) within-participants design. The dependent variable was the recognition decision for each test word.

Materials

The materials for this experiment were subdivided into a conjunction lure set and a semantic lure set. Each set of items consisted of 405 words forming 135 triplets. In the conjunction lure set each triplet consisted of two parent compound words (such as ''tailspin'' and ''floodgate'') and a conjunction lure created by recombining the syllables of the parent words (such as ''tailgate''). Efforts were taken to ensure that the semantic overlap between the parent words and the conjunction lures was minimised. For example, parent/lure combinations in which one morpheme was used in the same sense, such as ''storm'' in ''thunderstorm'' and ''hailstorm'', were excluded from the list. In the semantic lure set each triplet consisted of three words that were synonyms of one another (such as ''fury'', ''anger'', and ''rage''). For this set the semantic overlap between the items in each triplet was maximised. The triplets were primarily based on normed lists of synonyms (Whitten, Suter & Frank, 1979; Wilding & Mohindra, 1981, 1983) and included sets of words that were rated as being highly similar in meaning.

We used two parent items in both the conjunction and semantic lure sets for two reasons: to keep the structures of the experimental lists identical for the two conditions and to increase the number of false alarms in both conditions. Past experiments using conjunction lures have found that while participants false alarm to lures that overlap with only one parent item (feature lures), the false alarm rates are higher when the lures overlap with two parent items (Reinitz et al., 1992; Underwood et al., 1976; Underwood & Zimmerman, 1973). For semantic lures, false alarms are typically quite low for lures whose parents were studied in a list context (Matzen & Benjamin, 2009), but the false alarm rate increases when the study list contains multiple words that are semantically related to the lure (Hintzman, 1988; Roediger & McDermott, 1995; Shiffrin et al., 1995). Based on this past research we chose to use two parent items for both sets of lures.

Three counterbalancing conditions were used to divide the conjunction items into three lists. In the first counterbalancing condition two parent words were studied and then both were tested as old, to-be-endorsed items. In the second counterbalancing condition two parent words were studied in ''forward order'' (e.g., for the lure ''tailgate'', the parent word ''tailspin'' preceded the parent word ''floodgate'' in the study list) and the to-be-rejected conjunction lure was tested. In the third counterbalancing condition two parent words were studied in ''backward order'' (where the parent word ''floodgate'' preceded the parent word ''tailspin'' in the study list) and the conjunction lure was tested. These two presentation orders were used to make the relationship between the parent items and the lures less noticeable to the participants.

Nine counterbalancing conditions were used to divide the semantic items into lists. In three of the counterbalancing conditions two synonyms were studied and then both were tested as old, to-be-endorsed items. The words within each triplet were rotated through positions so that each word appeared equally often as the first parent word and the second parent word. In the other six counterbalancing conditions two synonyms were studied and the third was tested as a to-be-rejected semantic lure. The words within each triplet were rotated through list positions so that each word appeared equally often as the first parent word, second parent word, and semantic lure.

Two sets of new, unrelated items were created and matched to the items in the conjunction lure and semantic lure sets. Each set of new items contained 90 words. For the conjunction lure set the new items were compound words that do not share morphemes with any of the parent words or lures. For the semantic lure set the new items were common nouns that are not synonyms of any of the parent words or lures. Both types of new items were matched in terms of length and frequency with the items from their respective sets of lures. In the conjunction lure set the average length of the words was 8.5 letters for the parent items/old test items, 8.4 letters for the lures, and 9.1 letters for the new items. The average frequency was 2.8 for the parent items/old test items, 3.4 for the lures, and 2.0 for the new items. In the semantic lure set the average length of the words was 5.8 letters for the parent items, old test items, and lures, and 6.8 letters for the new items. The average frequency was 45.3 for the parent items, old test items, and lures, and 43.3 for the new items (frequency data were taken from the Kucera & Francis, 1967; norms included in Balota et al., 2002; a frequency value of zero was assumed for items not appearing in the database).

The words in the conjunction lure and semantic lure sets were placed in a pseudorandom order with the appropriate parent items, old items, lures, and new items placed in the appropriate slots in each of the different counterbalanced lists. The participants studied parent words and were later presented with the same word or with a related lure at one of nine lags: 1, 2, 3, 5, 10, 25, 40, 60, or 100. The pairs of parent items that were related to the same lure were always presented back-to-back with no other items intervening between them. For each lag *N*, there were *N*–1 words between the second word in the pair of studied parent items and the corresponding test item. Across the nine experimental lists each item was rotated through all nine possible lags using a Latin square design. Each list had 15 pairs of parent items appearing in each lag condition, 10 of which were later tested with their corresponding lure word (leading to 10 lures for each lag condition) and 5 of which were later tested as old items (leading to 10 old items for each lag condition). The final experimental lists for each item set contained a total of 270 study words consisting of 180 parent words that were later tested with their corresponding lures and 90 parent words that were later tested as old items. Intermixed with the study words at the appropriate lags there were 270 test words consisting of 90 lures, 90 old items, and 90 new, unrelated items. Altogether, each experimental list contained 540 words.

During the experiment each participant was presented with one of the conjunction lure lists and one of the semantic lure lists with a short break in between. The order in which the lists were presented was counterbalanced across participants.

Procedure

Participants were instructed that they would see a list of words to study and that intermixed with the study words there would be test items that may or may not have appeared earlier in the list. All of the study words were presented in the centre of the computer screen in black 18-point Courier New font on a white background. Each study word was presented for 600 milliseconds with a 250 ms interstimulus interval. The test items appeared in red font and were interspersed with the study items. The test words remained on the computer screen until the participant made a response. Once the participant responded to a test word, that word disappeared from the screen. It was followed by a 250-ms ISI, then by another study or test word.

The participants were instructed that when they saw a red word they were to respond by pressing one of four keys on the keyboard. If they thought that they studied the word earlier in the list they were asked to press "R" if they specifically remembered seeing the word or "F" if they did not specifically remember it but thought that it seemed familiar. If they thought that they did not study the word earlier in the list they were asked to press the "U" key if the word simply seemed unfamiliar or the "D" key if they knew that the test word is new because they remembered seeing a similar but slightly different word earlier in the study list. The participants were given a sheet of paper explaining each of the four possible responses that they could refer to throughout the experiment.

Analysis

The rates of each type of response were calculated for each type of item (old, lure, and new conjunction items and old, lure, and new semantic items) at each lag. The "Different" responses were used as a measure of recollection rejection while the "Remember" and "Familiar" responses were used to plot the false alarm functions for each lure type.

RESULTS

The mean proportions of each response type were calculated for each item type at all of the nine lags. These are shown in Table 1. The participants'

TABLE 1
Mean proportions of each response type at each lag for all items

		Old items				Lures				New items			
	Lag	Remember	Familiar	Unfamiliar	Different	Remember	Familiar	Unfamiliar	Different	Remember	Familiar	Unfamiliar	Different
Conjunction item set	1	0.92	0.04	0.01	0.03	0.36	0.09	0.17	0.38				
	2	0.68	0.14	0.11	0.06	0.49	0.16	0.16	0.19				
	3	0.67	0.14	0.11	0.08	0.50	0.17	0.17	0.16				
	5	0.44	0.19	0.25	0.12	0.37	0.24	0.26	0.13				
	10	0.49	0.22	0.21	0.08	0.27	0.23	0.34	0.16	0.08	0.08	0.64	0.20
	25	0.43	0.19	0.26	0.12	0.23	0.21	0.40	0.16				
	40	0.33	0.21	0.35	0.11	0.26	0.22	0.38	0.15				
	60	0.41	0.22	0.24	0.12	0.19	0.22	0.42	0.18				
	100	0.30	0.19	0.35	0.16	0.17	0.19	0.45	0.19				
Semantic item set	1	0.95	0.02	0.01	0.02	0.10	0.09	0.32	0.48				
	2	0.77	0.09	0.09	0.04	0.13	0.11	0.38	0.38				
	3	0.79	0.09	0.08	0.04	0.13	0.16	0.38	0.33				
	5	0.56	0.21	0.16	0.08	0.19	0.14	0.41	0.26				
	10	0.51	0.20	0.18	0.11	0.16	0.18	0.40	0.26	0.11	0.14	0.57	0.18
	25	0.48	0.19	0.22	0.11	0.14	0.14	0.49	0.23				
	40	0.46	0.21	0.21	0.12	0.11	0.16	0.49	0.25				
	60	0.42	0.22	0.23	0.13	0.13	0.15	0.48	0.24				
	100	0.40	0.21	0.25	0.15	0.15	0.14	0.49	0.22				

responses to the old items in both the conjunction lure set and the semantic lure set were quite similar. *F*-tests conducted on these data revealed that the responses to the old items did not differ significantly across item sets (all $Fs < 1.93$, all $ps > .19$), so these data were collapsed in subsequent analyses.

Looking at the collapsed data for the old test items, the proportion of "remember" responses decreased as lag increased and the proportions of other three response types increased as lag increased, as expected. A large majority of participants (98%) showed a decrease in "remember" responses from lag 1 to lag 100 and an increase in "different" (72%), "familiar" (81%), and unfamiliar (91%) responses. The average proportion of each response type at each lag for the combined old items is shown in Figure 1.

Our lure-type analysis rests on the assumption that the two types of lures would exhibit non-monotonic false alarm functions. The relevant data are shown in Figure 2. As discussed above, our primary interest is in the trade-offs between familiarity and recollection in participants' decision making about these two types of lures. The crucial information for addressing this issue comes from comparing of the locations of the peaks of the two false alarm functions. Critically, we predicted that the point at which the false alarm functions curved would be different for the

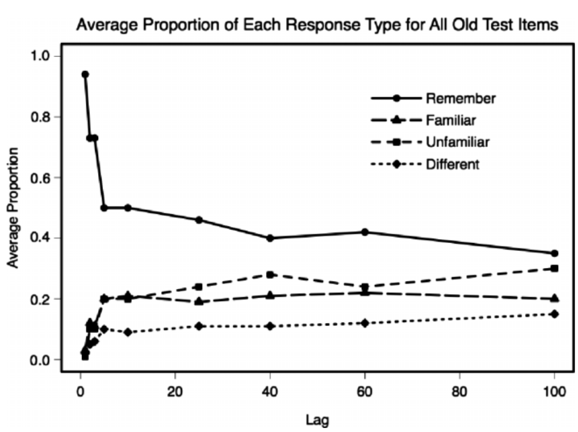

Figure 1. The average proportion of each response type for the old items presented at test. When participants judged a test item to be old, they specified whether they specifically remembered seeing that item at study ("Remember") or whether they did not specifically remember seeing the item but still thought that it has been on the study list ("Familiar"). When participants judged a test item to be new, they specified whether they were certain that it was new because they specifically remembered studying a similar word ("Different") or whether they simply thought that the word had not been on the study list ("Unfamiliar").

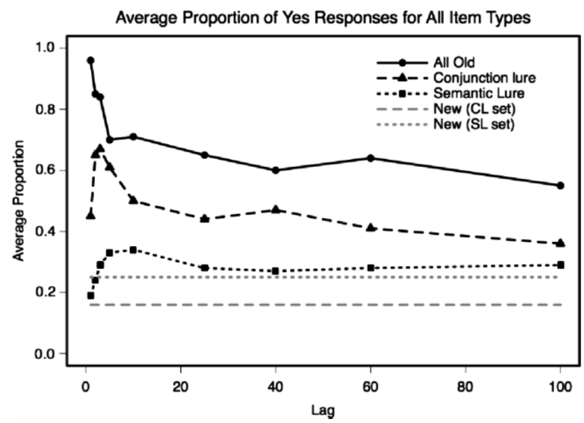

Figure 2. The average proportion of "Yes" responses for all item types. "Yes" responses represent the combination of the "Remember" and "Familiar" responses. The item types were conjunction lures, semantic lures, new items from the conjunction and semantic lures sets, and the combined old items from the conjunction and semantic lures sets.

two types of lures, with the non-monotonicity occurring at a shorter lag for the conjunction lures. In order to plot the false alarm functions, the false alarms for each lure type were calculated by combining both types of "yes" responses ("remember" and "familiar") at each lag. As predicted, the data showed that the false alarm functions were non-monotonic for both lure types. Additionally, for the group averages, false alarm rates peaked for the conjunction lures at lag 3 while the false alarm rate for the semantic lures peaked later, at lag 10.

To test this difference in peaks across individuals, we recorded for each participant the lag (or lags) where the false alarms were at a maximum, both for semantic lures and conjunction lures. If a participant had multiple maxima, the average of the lags was recorded. Across participants, the difference max_{CL}–max_{SL} was significantly greater than zero, $t(57) = 4.17$, $p < .01$.

Our final prediction was that there would be higher rates of recollection rejection for semantic lures than for conjunction lures. To test this prediction we calculated the average proportion of "different" responses for each item type, which indicated that the participants recalled studying a word that was similar but different from the test word. These data are shown in Figure 3. Note that the overall pattern of recollection rejection responses across lags is quite similar to what Jones and Atchley (2006) found for conjunction lures. As predicted, the proportion of "different" responses was higher for semantic lures than for conjunction lures, with the difference between the

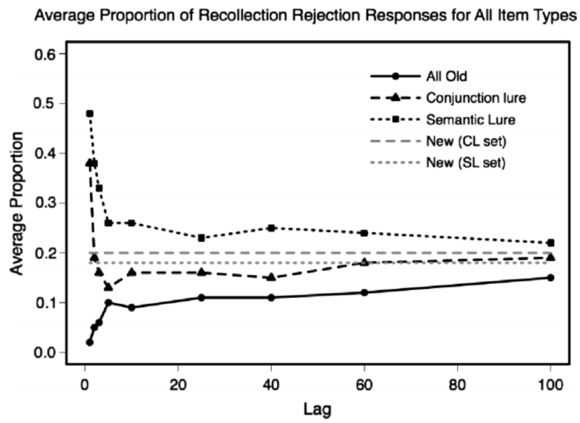

Figure 3. The average proportions of recollection rejection responses for all item types. The proportion of ‘‘Different’’ responses for each item type is used as a measure of recollection rejection.

two significant at all lags except for lag 100—all ts > 2.18; all ps < .03 except for $t(57) = 1.19$ at lag 100. For conjunction lures the rate at which participants chose the ‘‘different’’ response dropped down to the base rate of that response for unrelated new items by lag 3. The rate at which participants chose the ‘‘different’’ response for semantic lures was higher than the rate for unrelated new items from the semantic set at all lags (all ts > 2.06; all ps < .04).

The results of this experiment revealed that the false alarm functions peak at different times for semantic and conjunction lures. Additionally, the analysis of the ‘‘different’’ response indicated a higher rate of recollection rejection for semantic lures than for conjunction lures across all lags. While these data show that different rates of recollection rejection contribute to the differences in the false alarm functions for semantic and conjunction lures, the contributions of familiarity are less clear. We used a multinomial process model, described in the following section, to tease apart the contributions of familiarity and recollection to each response type.

Multinomial process model

To better understand the roles of familiarity and recollection rejection in recognition for semantic and conjunction lures, we constructed a multinomial process model representing our assumptions about these two processes. The goal of the model was to trace the degree of familiarity and the probability of recollection rejection for old items, new items, and lures across the lags. The advantage of this analysis over the analyses based entirely on individual response types is that the hypothetical latent variables—familiarity and recollection—bear different relationships with responses across conditions, and the previous analysis does not use these principles to combine results across response types in a theoretically motivated way. To our knowledge, this is the first model to fit decay functions to surface-level and semantic information decay, providing a new look at the interplay between recollection and familiarity.

A schematic of the model is portrayed in Figure 4. To begin, we assume that each lure item either does or does not elicit sufficient familiarity to encourage assessment of the item as previously seen. Each item also may or may not successfully remind the participant of a previous study event; if that event contained the stimulus being probed, that recollection promotes endorsement of the item. If that event involved a similar but nonetheless different stimulus, that recollection becomes recollection rejection, and the item is rejected despite its familiarity. Lure items that are familiar but do not lead to a reminding result in either a Remember or Familiar response, depending on the value of C, a response bias parameter. The same bias parameter is used in all four models, reflecting our assumption that the tendency to say Remember vs Familiar does not depend on the item type.

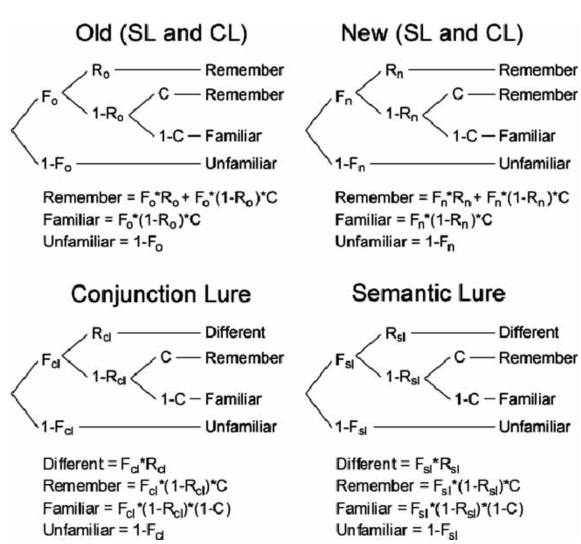

Figure 4. A schematic representation of the decision trees for each item type that were used in the multinomial process model.

Items that do not elicit sufficient familiarity result in the Unfamiliar response.[1]

Note that we are assuming that the Remember and Familiar responses do not reflect qualitatively different processes, but rather different response criteria. These two responses could also be interpreted as reflecting different states of knowledge, but in either case there is some set of information that pushes participants towards one response or the other, and the model does not depend on how the difference between Remember and Familiar is specifically defined.

As for any model, ours is a simplification of the steps engaged in to reach a response, but it does capture the two central processes that are likely attempted on each trial—familiarity assessment and recollection. Other response pathways are technically possible (e.g., participants idiosyncratically perceive a new item as abstractly related to but Different from a studied item), but those pathways are not our focus here. Model fitting is crucial for interpreting memory performance, because responses are rarely ‘‘process pure.’’ By plotting the model fits over time, we have the unique opportunity to assess the differential roles of familiarity and recollection in yielding the differential peaks for conjunction lures and semantic lures shown in Figure 2.

We fitted the models separately for each participant and each lag. Fits were obtained through maximum likelihood estimation, with likelihoods generated from the tree diagrams in Figure 4. For example, the probabilities of responding Different, Remember, Familiar, and Unfamiliar for a conjunction lure item were $F_{cl}R_{cl}$, $F_{cl}(1–R_{cl})C$, $F_{cl}(1–R_{cl})(1–C)$, and $1–F_{cl}$, respectively, and the likelihood for response distribution $[d, r, f, u]$ was $d^{F_{cl}R_{cl}} r^{F_{cl}(1-R_{cl})C} f^{F_{cl}(1-R_{cl})(1-C)} u^{(1-F_{cl})}$. The overall likelihood for the joint multinomial model was the product of six item-specific likelihoods—that of the Old (SL set) items, Old (CL set), New (SL set), New (CL set), Semantic lures, and Conjunction lures. Note that two separate models were fitted to the semantic lure (SL) set and the conjunction lure (CL) set for old and new items. This was done so that we may observe any crucial differences in the model fits across the item sets. (To preview, the differences were small.)

The fitted familiarity and reminding parameters are plotted in Figure 5. Each data point represents the mean parameter value, across participants, for a given lag. Before exploring the fits any further, we obtained the model’s goodness of fit using likelihood ratio tests. These tests compared, for each participant, the fits of our model to the fits of the general multinomial model—a model that assigns a unique parameter to each datum, perfectly fitting the data. The model adequately fits the data if its performance is comparable to the general multinomial model, or in other words if the ratio of likelihoods is

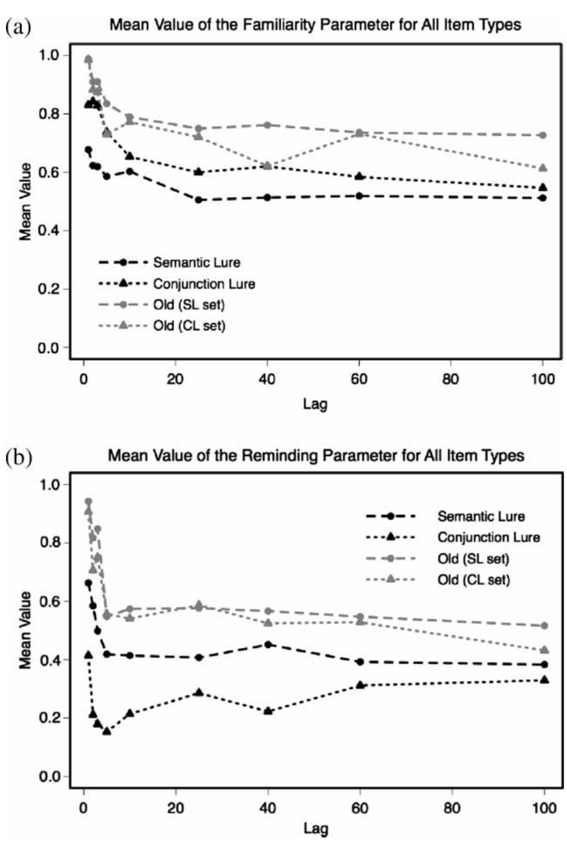

Figure 5. Mean familiarity (a) and reminding (b) parameter values across participants, per lab and item type. In both figures the old items are divided into those from the semantic lure set and those from the conjunction lure set.

[1] Although familiarity assessment is the first/leftmost process in the tree diagram for our model, it is also possible to construct a model where recollection is first. This model is identical to our model for items that are both familiar and not recollected. For other items there are three differences in the recollection first model. First, the lures that are recollected lead to a ‘‘Different’’ response regardless of familiarity. Second, the old and new items that are recollected lead to a ‘‘Remember’’ response regardless of familiarity. Third, items elicit an ‘‘Unfamiliar’’ response when they are neither familiar nor recollected. We compared the likelihood of this model to our original model and found that our original model was greater for 48 of 58 participants. Of the 10 participants who fitted better with the recollection first model, only 4 had an AIC with a greater than 5-point disparity from our original model; hence 54 of 58 participants were not fitted significantly less well by our original model.

small (Riefer & Batchelder, 1988). For over 97% of participants, the model's fit was not significantly worse than that of the general multinomial model—i.e., for over 97% of participants, $\chi^2(6) < 16.81$; $\alpha = .01$—indicating good overall fits. The goodness of fit was comparable (in fact, slightly better than) to that reported in similar applications of multinomial models to memory data (e.g., McBride, Dosher, & Gage, 2001). We also computed AIC and BIC (Akaike and Bayesian Information Criteria, respectively) for the general model and our model. The AIC of our model was lower (indicating a better fit) for 97% of participants. BIC, which greatly prefers simpler models, was lower for 100% of participants. Hence, the sparse but theoretically motivated response pathways were sufficient to explain subjects' performance.

To test our hypothesis regarding the differential decay of familiarity for semantic lures and conjunction lures, we fit participant's familiarity parameters across the nine lags to a power function ($y = b \cdot t^{-d}$) using hierarchical logistic regression. The regression equation represents the log-odds of participants' F parameters as a function of log(lag), lure type, an interaction between log(lag) and lure type, and a random effect of participant, U_j, where $U_j \sim N(0,\tau^2)$.[2]

$$\log(\mathrm{F}_{ij}/(1-\mathrm{F}_{ij})) = \gamma_0 + \gamma_1 \log(lag_{ij}) + \gamma_2 luretype_{ij} + \gamma_3 \log(lag) luretype_{ij} + U_j \quad (1)$$

Taking the exponential of both sides of the equation shows that this model relates the odds of F (i.e., the odds of the test item being familiar) to a power function of lag (see Appendix for algebraic derivation):[3]

$$F_{ij}/(1-F_{ij}) = e^{\gamma_0 + \gamma_2 luretype_{ij} + U_j} lag_{ij}^{\gamma_1 + \gamma_2 luretype_{ij}} \quad (2)$$

R-squared values for the conjunction lure and semantic lure regressions were 0.98 and 0.97, respectively, indicating good overall fits. Visual inspection of Figure 5a and 5b also confirms that the regression fits closely follow the patterns in the data. Wald tests on parameter fits showed a significant effect of lag, $z^2(983) = 2.76$, $p < .05$, a significant effect of lure type, $z^2(983) = 4.16$, $p < .05$, and a significant interaction between lag and lure type, $z^2(983) = 2.29$, $p < .05$. A further test of the interaction showed that a model with no interaction term fitted worse (i.e., the likelihood was lower) than the model including the interaction, $\chi^2(1) = 0.49$, $p < .05$. The regression with the interaction term indicated that the rate of decay of familiarity was 2.25 times greater for conjunction lures than for semantic lures (95% confidence interval: $1.58 < x < 2.38$). Hence, our analysis does reveal that surface information decays faster than semantic information.

To test the second hypothesis regarding reminding leading to recollection rejection, we compared the reminding parameters for semantic lure and conjunction lure items at each lag using independent t-tests. Taking the average parameter value per participant across the nine lags, the R parameter was significantly greater for semantic lures than conjunction lures, $t(57) = 7.54$, $p < .05$, With a corrected alpha for multiple comparisons (9 lags; alpha = 0.05/9 = 0.006), the R parameter was significantly greater for semantic lures than conjunction lures at lags 1, 2, 3, 5, 10, and 40, providing evidence that reminding was more likely for semantic lures than conjunction lures, at least at these early lags.

DISCUSSION

The behavioural data from this experiment show that, as predicted, the peak of the non-monotonic false alarm function occurs at longer study–test lags for semantic lures than for conjunction lures. While the false alarm rate for conjunction lures peaked at lag 3 and declined across subsequent lags, the false alarm rate for semantic lures did

---

[2] The random effect of participants, U_j, captures the deviation of individuals from the group-level intercept, γ_0. The model assumes that individuals sampled in our study come from an underlying distribution with mean 0 and variance τ^2, a free parameter of the regression. This enables us to generalise to that distribution, rather than to the particular set of participants in our study, as a fixed-effects model (e.g., ANCOVA) would restrict us to (see Snijders & Bosker, 1999). The τ^2 parameter also helps to represent the possible participant-level error in estimating the parameters of the multinomial model.

[3] We chose the power function due to usage in previous work comparing semantic and word-form information decay (e.g., McBride et al., 2001) and based on the outcome of power and exponential fits to individual participants using MLE. Of the fits for which both power and exponential models converged (see Myung, 2003, for MLE settings), 67% received a higher likelihood with the power function. Furthermore, only 13% of the power fits had an AIC with a greater than 5 point disparity from the exponential fit, suggesting the power function did a good job of describing the majority of participants.

not peak until lag 10. To our knowledge this is the first experiment to compare the false alarm functions for these two types of lures. However, past research suggests two hypotheses for why the peak of the non-monotonicity should occur at longer lags for lures that are semantically related to studied items as opposed to lures that share only a surface similarity with studied items.

First, we hypothesised that semantic lures would be more likely to remind participants of previously studied words, leading to recollection rejection. The "Different" and "Unfamiliar" judgements that the participants made for all "no" responses in the recognition test were used to test this hypothesis. These data show that the participants were more likely to use recollection rejection for semantic lures than for conjunction lures across all lags. This is consistent with previous research suggesting that semantic relationships are more likely to lead to recollection rejection than surface form overlap (Odegard & Lampinen, 2005; Odegard et al., 2005) and supports our prediction that different rates of recollection rejection would shift the peak of the false alarm function for semantic lures to longer lags relative to the peak of the false alarm function for conjunction lures.

Our second prediction was that there would be a more rapid loss of familiarity for surface form information than for semantic information as the lag between study and test increased. If, as suggested by previous research on word and sentence processing (Bock & Brewer, 1974; Brewer, 1975; Matzen & Benjamin, 2009; Potter & Lombardi, 1990; Sachs, 1967), semantic information is encoded more robustly than surface form information, the surface form information will become inaccessible more rapidly as the time between study and test increases and the overall memory trace degrades (Benjamin, 2008). As the surface form information becomes less accessible, the conjunction lures become less and less likely to seem familiar. The multinomial process model demonstrates that, as predicted, familiarity decreases more rapidly for conjunction lures than for semantic lures. The model itself is the first to examine recollection and familiarity separately by fitting decay functions to surface-level and semantic information decay. It supports our hypothesis that a differential loss of familiarity for the two types of lures is another factor that would shift the peaks of their false alarm functions relative to one another.

We note that our model is one of many that have been used to understand the role of familiarity and recollection in recognition judgements. Our model depicts the participants' task in this experiment as one similar to many real-world cases in which people first feel a sense of familiarity and then try to account for that feeling through recollection of details. Although other dual-process models assume that familiarity is secondary to recollection (e.g., Brainerd et al., 2003; McBride & Dosher, 2001; Yonelinas, 1999), our model claims the opposite—that recollection is used secondarily to "account for" or "follow up" positive familiarity assessments. This design choice was validated empirically by fitting a separate model with recollection first (and showing a worse fit; see footnote 1), and is supported by evidence showing that familiarity is faster to assess than recollection (Benjamin & Bjork, 2000; Hintzman & Curran, 1994). We find this outcome interesting, but do *not* take it to mean that previous models were wrong. The two processes may be prioritised differently according to the goals of the task at hand. The influence of goals and task factors on which process takes priority is a topic that merits additional research.

This study reveals the interaction of two processes that differentially affect memory for surface form and semantic information. The first of these processes is familiarity. As the time between study and test increases, the familiarity elicited by overlapping surface forms across study and test words decreases faster than the familiarity elicited by semantic overlap. This could be because readers encode less information about surface form to begin with, or because the smaller set from which surface form information is drawn causes interference and leads people to change their response criteria (Benjamin & Bawa, 2004). Surface form and semantic information also differ in their ability to serve as effective cues for recollection rejection. A lure that shares semantic features with its parent word is more likely to trigger recollection of that parent word than a lure that shares only surface features with its parent. This difference in rates of recollection rejection and the more rapid decrease in familiarity for surface form information combine to shift the peak of the false alarm rate function to later study–test lags for semantic lures relative to conjunction lures.

The results of this experiment also provide an explanation for why many studies of conjunction lures have not found any evidence of recollection rejection (Jones, Brown, & Atchley, 2007; Odegard & Lampinen, 2005; Odegard et al., 2005; Wong & Rotello, 2010). While some experiments have found recollection rejection for conjunction lures only when the lures share semantic features with their parent items (Odegard & Lampinen, 2005; Odegard et al., 2005), others have failed to find evidence of recollection rejection, even with lures that have substantial semantic overlap with their parent items (Wong & Rotello, 2010). One explanation for these discrepancies could be the lag between the studied items and the test items, both in terms of the distance between the parent items in the study list and in terms of the time between the study and the test list. Our study, like Jones and Atchley (2006), found evidence of recollection rejection for conjunction lures that did not share semantic features with their parents only at very short lags (in the present experiment this occurred only when the lure immediately followed the parent items). However, even for semantic lures, rates of recollection rejection dropped of substantially as the lag between study and test increased. Conjunction lures that share semantic overlap with their parent items, such as those used in Wong and Rotello (2010) have both semantic and surface form overlap with their parent items. These lures may be more difficult to reject using recollection rejection, because the participants would need to recall a great deal of detail about the originally studied words in order to reject the very similar lures. As the time between study and test increases, this would become even more difficult. The lures that share semantic features with their parents will still seem familiar (which is consistent with Wong and Rotello's finding of higher false alarm rates for lures that were semantically transparent), but perhaps too much information about the surface forms of the studied items has been lost at longer lags, so participants either do not use a recollection rejection strategy or that strategy is unsuccessful.

An interesting topic for future research would be mapping the false alarm function for conjunction lures that share semantic features with their parent items. One might predict that the peak of that function would fall between the peaks of the two functions mapped in the present experiment. The greater semantic overlap would lead to a greater sense of familiarity for the lures, relative to conjunction lures that do not share semantic features with their parent items, but it would also support higher rates of recollection rejection. At the same time the high degree of surface form overlap could make recollection rejection more difficult, so rates of recollection rejection may be lower than they would be for synonym lures that have no surface form overlap.

In a broader sense, the results of this study show how recognition memory errors in response to lures with specific relationships to their parent items can reveal a great deal about the processes underlying memory performance for different kinds of information. Comparing different kinds of lures to one another, especially as a function of time between study and test, can be very informative as to how familiarity, recollection, and forgetting affect different kinds of encoded information and lead to different patterns of both true and false memory.

Manuscript received 17 March 2010
Manuscript accepted 28 September 2010

REFERENCES

Balota, D. A., Cortese, M. J., Hutchison, K. A., Neely, J. H., Nelson, D., Simpson, G. B., et al. (2002). *The English Lexicon Project: A web-based repository of descriptive and behavioral measures for 40,481 English words and nonwords.* Washington University. http://elexicon.wustl.edu/

Bellezza, F. S., Winkler, H. B., & Andrasik, F. (1975). Encoding processes and the spacing effect. *Memory & Cognition*, *3*, 451–457.

Benjamin, A. S. (2001). On the dual effects of repetition on false recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *27*, 941–947.

Benjamin, A. S. (2005). Recognition memory and introspective remember/know judgements: Evidence for the influence of distractor plausibility on "remembering" and a caution about purportedly nonparametric measures. *Memory and Cognition*, *33*, 261–269.

Benjamin, A. S. (2008). Memory is more than just remembering: Strategic control of encoding, accessing memory, and making decisions. In A. S. Benjamin & B. H. Ross (Eds.), *The psychology of learning and motivation: Skill and strategy in memory use* (Vol. 48, pp. 175–224). London: Academic Press.

Benjamin, A. S., & Bawa, S. (2004). Distractor plausibility and criterion placement in recognition. *Journal of Memory & Language*, *51*, 159–172.

Benjamin, A. S., & Bjork, R. A. (2000). On the relationship between recognition speed and accuracy for words rehearsed via rote versus elaborative rehearsal. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *26*, 638–648.

Bock, J. K., & Brewer, W. F. (1974). Reconstructive recall in sentences with alternative surface structures. *Journal of Experimental Psychology*, *103*(5), 837–843.

Brainerd, C. J., & Reyna, V. F. (1998). Fuzzy trace theory and children's false memories. *Journal of Experimental Child Psychology*, *71*, 81–129.

Brainerd, C. J., & Reyna, V. F. (2001). Fuzzy-trace theory: Dual processes in memory, reasoning, and cognitive neuroscience. *Advances in Child Development and Behavior*, *6*, 359–364.

Brainerd, C. J., & Reyna, V. F. (2007). Explaining developmental reversals in false memory. *Psychological Science*, *18*, 442–448.

Brainerd, C. J., Reyna, V. F., Wright, R., & Mojardin, A. H. (2003). Recollection rejection: False-memory editing in children and adults. *Psychological Review*, *110*, 762–784.

Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology*, *2*, 331–350.

Bray, J. F., & Robbins, D. (1976). Frequency judgements and the spacing effect: Immediate and delayed performance. *Bulletin of the Psychonomic Society*, *7*, 47–49.

Brewer, W. F. (1975). Memory for ideas: Synonym substitution. *Memory and Cognition*, *3*(4), 458–464.

Brewer, W. F. (1977). Memory for the pragmatic implications of sentences. *Memory and Cognition*, *5*(6), 673–678.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*, 17–22.

Dunn, J. C. (2004). Remember–know: A matter of confidence. *Psychological Review*, *111*, 524–542.

Gallo, D. A. (2006). *Associative illusions of memory*. New York: Psychology Press.

Gallo, D. A., Meadow, N. G., Johnson, E. L., & Foster, K. T. (2008). Deep levels of processing elicit a distinctiveness heuristic: Evidence from the criterial recollection task. *Journal of Memory and Language*, *58*, 1095–1111.

Greene, R. L. (2004). Recognition memory for pseudowords. *Journal of Memory and Language*, *50*, 259–267.

Heathcote, A. (2003). Item recognition memory and the receiver operating characteristic. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 1210–1230.

Hintzman, D. L. (1976). Repetition and memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 10, pp. 47–91). New York: Academic Press.

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

Hintzman, D. L., & Curran, T. (1994). Retrieval dynamics of recognition and frequency judgements: Evidence for separate processes of familiarity and recall. *Journal of Memory & Language*, *33*, 1–18.

Jacoby, L. L. (1996). Dissociating automatic and consciously controlled effects of study/test compatibility. *Journal of Memory and Language*, *35*, 32–52.

Jacoby, L. L., & Craik, F. I. M. (1979). Effects of elaboration of processing at encoding and retrieval: Trace distinctiveness and recovery of initial context. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 1–21). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Johnson, M. K., Bransford, J. D., & Solomon, S. K. (1973). Memory for tacit implications in sentences. *Journal of Experimental Psychology*, *98*, 203–205.

Jones, T. C. (2005). Study repetition and the rejection of conjunction lures. *Memory*, *13*, 499–515.

Jones, T. C., & Atchley, P. (2002). Conjunction error rates on a continuous recognition memory test: Little evidence for recollection. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 374–379.

Jones, T. C., & Atchley, P. (2006). Conjunction errors, recollection-based rejections, and forgetting in a continuous recognition task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 70–78.

Jones, T. C., Brown, A. S., & Atchley, P. (2007). Feature and conjunction effects in recognition memory: Toward specifying familiarity for compound words. *Memory & Cognition*, *35*, 984–998.

Jones, T. C., & Jacoby, L. L. (2001). Feature and conjunction errors in recognition memory: Evidence for dual-process theory. *Journal of Memory and Language*, *45*, 82–102.

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

Lampinen, J. M., Odegard, T. N., & Neuschatz, J. S. (2004). Robust recollection rejection in the memory conjunction paradigm. *Journal of Experimental Psychology: Learning Memory & Cognition*, *30*, 332–342.

Lloyd, M. E. (2007). Metamemorial influences in recognition memory: Pictorial encoding reduces conjunction errors. *Memory & Cognition*, *35*, 1067–1073.

MacLeod, C. M. & Nelson, T. O. (1976). A non-monotonic lag function for false alarms to associates. *American Journal of Psychology*, *89*, 127–135.

Matzen, L. E., & Benjamin, A. S. (2009). Remembering words not presented in sentences: How study context changes patterns of false memories. *Memory & Cognition*, *37*, 52–64.

Marsh, R. L., Hicks, J. L., & Davis, T. T. (2002). Source monitoring does not alleviate (and may exacerbate) the occurrence of memory conjunction errors. *Journal of Memory and Language*, *47*, 315–326.

McBride, D. M., Dosher, B., & Gage, N. (2001). A comparison of forgetting for conscious and automatic memory processes in word fragment completion tasks. *Journal of Memory and Language*, *45*, 585–615.

McDermott, K. B., & Watson, J. M. (2001). The rise and fall of false recall: The impact of presentation duration. *Journal of Memory and Language*, *45*, 160–176.

Miller, M.B., & Wolford, G. L. (1999). The role of criterion shift in false memory. *Psychological Review*, *106*, 398–405.

Moscovitch, M., & Craik, F. I. M. (1976). Depth of processing, retrieval cues, and uniqueness of encoding as factors in recall. *Journal of Verbal Learning and Verbal Behavior*, *15*, 447–458.

Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, *47*, 90–100.

Odegard, T. N., & Lampinen, J. M. (2005). Recollection rejection: Gist cuing of verbatim memory. *Memory & Cognition*, *33*, 1422–1430.

Odegard, T. N., Lampinen, J. M., & Toglia, M. P. (2005). Meaning's moderating effect on recollection rejection. *Journal of Memory and Language*, *53*, 416–429.

Potter, M. C., & Lombardi, L. (1990). Regeneration in the short-term recall of sentences. *Journal of Memory and Language*, *29*, 633–654.

Ratcliff, R. & Hockley, W. E. (1980). Repeated negatives in item recognition: Non-monotonic lag functions. In R. S. Nickerson (Ed.), *Attention and performance VIII*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Reinitz, M. T., Lammers, W. J., & Cochran, B. P. (1992). Memory-conjunction errors: Miscombination of stored stimulus features can produce illusions of memory. *Memory and Cognition*, *20*(1), 1–11.

Riefer, D. M., & Batchelder, W. H. (1988). Multinomial modelling and the measurement of cognitive processes. *Psychological Review*, *95*, 318–339.

Roediger, H. L. (1996). Memory illusions. *Journal of Memory and Language*, *35*, 76–100.

Roediger, H. L., & McDermott, K. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *21*, 803–814.

Roediger, H. L. III, McDermott, K. B., & Robinson, K. J. (1998). The role of associative processes in creating false memories. In M. A. Conway, S. E. Gathercole, & C. Cornoldi (Eds.), *Theories of memory II* (pp. 187–245). Hove, UK: Psychological Press.

Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception & Psychophysics*, *2*, 437–442.

Schacter, D. L., Israel, L., & Racine, C. (1999). Suppressing false recognition in younger and older adults: The distinctiveness heuristic. *Journal of Memory and Language*, *40*, 1–24.

Shiffrin, R. M., Huber, D. E., & Marinelli, K. (1995). Effects of category length and strength on familiarity in recognition. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *21*, 267–287.

Snijders, T., & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modelling*. Thousand Oaks, CA: Sage.

Thios, S. J., & D'Agostino, P. R. (1976). Effects of repetition as a function of study-phase retrieval. *Journal of Verbal Learning and Verbal Behavior*, *15*, 529–537.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist*, *26*, 1–12.

Underwood, B. J., Kapelak, S. M., & Malmi, R. A. (1976). Integration of discrete verbal units in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory*, *2*, 293–300.

Underwood, B. J., & Zimmerman, J. (1973). The syllable as a source of error in multisyllable word recognition. *Journal of Verbal Learning and Verbal Behavior*, *12*, 701–706.

Whitten, W. B., Suter, W. N., & Frank, M. L. (1979). Bidirectional synonym ratings of 464 noun pairs. *Journal of Verbal Learning and Verbal Behavior*, *18*, 109–127.

Wilding, J., & Mohindra, N. (1981). Ratings of the degree of synonymity of 279 noun pairs. *British Journal of Psychology*, *72*, 231–240.

Wilding, J., & Mohindra, N. (1983). Preferred synonyms for each noun of 279 noun pairs. *British Journal of Psychology*, *74*, 91–100.

Wixted, J. T. (2009). Remember/Know judgements in cognitive neuroscience: An illustration of the underrepresented point of view. *Learning and Memory*, *16*, 406–412.

Wong, M., & Rotello, C. M. (2010). Conjunction errors and semantic transparency. *Memory & Cognition*, *38*, 47–56.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*, 441–517.

APPENDIX

Our linear regression equation is the following:

$$\log(F_{ij}/(1-F_{ij})) = \gamma_0 + \gamma_1\log(lag_{ij}) + \gamma_2 luretype_{ij} + \gamma_3\log(lag)luretype_{ij} + U_j$$

Rearranging the sum on the right-hand side and then taking the exponential of both sides yields:

$$F_{ij}/(1-F_{ij}) = e^{\gamma_0+\gamma_2 luretype_{ij}+U_j+\gamma_1\log(lag_{ij})+\gamma_3\log(lag)luretype_{ij}},$$

which is equivalent to

$$F_{ij}/(1-F_{ij}) = e^{\gamma_0+\gamma_2 luretype_{ij}+U_j}e^{\gamma_1\log(lag_{ij})+\gamma_3\log(lag)luretype_{ij}}.$$

The terms in the second exponential can be rearranged to show a product of log(lag), as follows:

$$F_{ij}/(1-F_{ij}) = e^{\gamma_0+\gamma_2 luretype_{ij}+U_j}e^{\log(lag_{ij})[\gamma_1+\gamma_3 luretype_{ij}]},$$

which reduces to

$$F_{ij}/(1-F_{ij}) = e^{\gamma_0+\gamma_2 luretype_{ij}+U_j}lag_{ij}^{\gamma_1+\gamma_3 luretype_{ij}},$$

thus implying that the odds of an item being familiar, F_{ij}, are a power function of lag with base $e^{\gamma_0+\gamma_2 luretype_{ij}+U_j}$ and power $\gamma_1+\gamma_3 luretype_{ij}$.