MEMORY, 2002, *10* (4), 291–301

The breadth of memory search

Doug Rohrer
University of South Florida, USA

The recall of previously studied items is widely believed to incorporate a search of a markedly constrained set of possibilities, and the present study examines whether this set of items typically includes unstudied semantic associates of the study items. In an episodic task, participants recalled a previously studied list of eight exemplars drawn from a small or large category, and, in a semantic task, participants generated exemplars from these categories. Category size affected the time course of recall in the semantic task but not in the episodic task. This empirical dissociation between episodic and semantic memory is consistent with the view that episodic memory search efficiently excludes unstudied semantic associates of the study items and is instead constrained to those items sharing the temporal and spatial attributes of the episode.

INTRODUCTION

Perhaps the most salient characteristic of recall is the remarkably long pause that often precedes a response. When recalling a dozen words studied just 1 minute ago, for instance, subjects will typically recall at least one item after remaining silent for more than 10 seconds. Yet these lengthy pauses are arguably quite brief in light of the task at hand. The free recall of 12 unrelated words, for instance, requires one to choose from a lexicon of thousands of words. This feat is typically explained by assuming that memory retrieval is limited to a much smaller set of possibilities described here as the *search set*.

Although these search sets are widely believed to include study items that are not ultimately recalled, the present study examines whether these search sets typically include *unstudied* items. Notably, the presence of these so-called *extralist* items within the search set does not necessitate their overt recall, because extralist items might be rejected in a subsequent phase of retrieval. Of the various kinds of extralist items, perhaps the most likely candidates are semantic associates of the study items. Therefore, in order to maximise the likelihood of finding evidence that the search set can include extralist items, the present study required subjects to study and recall a list of words that comprised a subset of a semantic category. In one case, for example, subjects studied eight kinds of fruit, and these items were chosen randomly from a list including 24 exemplars. If extralist items are regularly included in the search set, at least some portion of the extralist ‘‘fruits’’ should be included as well.

Of further significance, the question of whether episodic retrieval relies on a search that regularly excludes extralist items is effectively a question about the distinction between episodic and semantic memory (cf. Tulving, 1986, 1993). That is, is the breadth of episodic search defined by the semantic attributes of the study items, or is episodic memory search limited to those items that share the temporal and spatial attributes of the episode itself? If the episodic retrieval of category exemplars requires a search through the entire set of both intralist and extralist category exemplars,

---

Requests for reprints should be sent to Doug Rohrer, Department of Psychology, PCD 4118G, University of South Florida, Tampa, FL 33620-7800, USA. Email: drohrer@chuma1.cas.usf.edu

I thank Kelli Taylor for her assistance with data collection; and Phil Beaman, Rick Hanley, Geoff Ward, and John Wixted for their comments and suggestions.

http://www.tandf.co.uk/journals/pp/09658211.html DOI:10.1080/09658210143000407

the search phase would not be easily distinguished from the kind of search underlying the generation of category exemplars from semantic memory. Hence, regardless of whether one is recalling eight previously studied kinds of fruits or simply generating kinds of fruit, this account would predict that both tasks rely on a search set including all kinds of fruit. By contrast, if episodic memory retrieval incorporates a search set that is limited to items from the episode, an attempt to recall eight previously studied fruits would incorporate a search that efficiently excludes the fruits that were excluded from the episode. By the latter account, episodic memory search would be truly episodic.

These theoretical issues are perhaps best illustrated by further description of the procedure. Participants were asked to study and freely recall a list of eight items drawn from either a small or large category in order to assess the effect of category size on recall. In particular, the eight-word study lists included virtually every member of a small category or only a fraction of the exemplars of a large category. Hence, this manipulation of category size effectively varied the number of extralist category exemplars. In turn, this manipulation of the number of extralist exemplars should affect search set size *if and only if* these extralist items are typically included in the search set. In order to provide the appropriate comparison, a second task required participants to generate category exemplars belonging to a given category. Category size is known to affect performance in this semantic task, as it naturally relies on a search of all category members. This experiment yields two plausible outcomes, and each is associated with one of the two competing hypotheses. If category size affects retrieval in both the episodic and semantic tasks, episodic retrieval presumably relies on a search set that includes the extralist category exemplars. Yet if category size affects retrieval in the semantic task but not in the episodic task, it will be argued that episodic retrieval relies on a search that typically excludes the extralist items.

The preceding rationale requires a behavioural measure of search set size, and the present study relies on a measure of latency. The use of a temporal measure is an intuitive choice, of course, because larger search sets presumably increase the time needed to find all of the ultimately recalled items. The validity of this measure is also demonstrated by several previous studies that are reviewed further later, after the following description of the measure itself.

Mean recall latency

The time course of recall is best illustrated by the distribution of recall latencies, and four such distributions are shown in Figure 1. These data are drawn from previous experiments that are described later, and the details of these studies are not critical here. In these plots, each data point represents the percentage of recalled words that were recalled in the corresponding 1-second interval. In the upper panel of Figure 1A, for example, about 13% of the responses were given in the second 1-second interval of the recall period. Latency distributions always exhibit an initial sharp ascent that reflects the 1–2-second pause that precedes the retrieval of the first item, and this pause includes non-memorial processes such as the perception of the category name (in the semantic task) or the instruction to begin recall (in the episodic task). Once retrieval begins, the rate of recall monotonically declines throughout the remainder of the recall period, as evidenced by the long tails of these distributions.

Mean recall latency equals the mean of the recall latency distribution. As is true with any distribution, the mean represents the location of the fulcrum upon which the distribution would balance, and these values are marked in Figure 1. Notably, mean recall latency can also be conceptualised as the arithmetic average of the latencies, regardless of whether the latencies are drawn from a single trial or from multiple trials. For example, if words are recalled at 5, 10, and 30 s into the recall period, mean recall latency equals 15 s [(5+10+30)/3]. Naturally, a brief mean recall latency entails that most of the responses were given early in the recall period, and a longer mean recall latency results when latencies are more evenly distributed throughout the recall period. Although not presented here, the time course of recall is also revealed by an analysis of inter-response times (IRTs) between consecutive responses (Patterson, Meltzer, & Mandler, 1971; Pollio, Richards, & Lucas, 1969; Rohrer, 1996; Rohrer & Wixted, 1994). IRT analyses are much more complex than latency analyses, however, because mean IRT must be determined as a function of both output position and recall total. Yet the two approaches yield the same conclusions (Rohrer, 1996), because the measures of latency and IRT are inherently dependent (McGill, 1963; Rohrer, 1996).

Although mean recall latency is easily calculated by simply determining the arithmetic

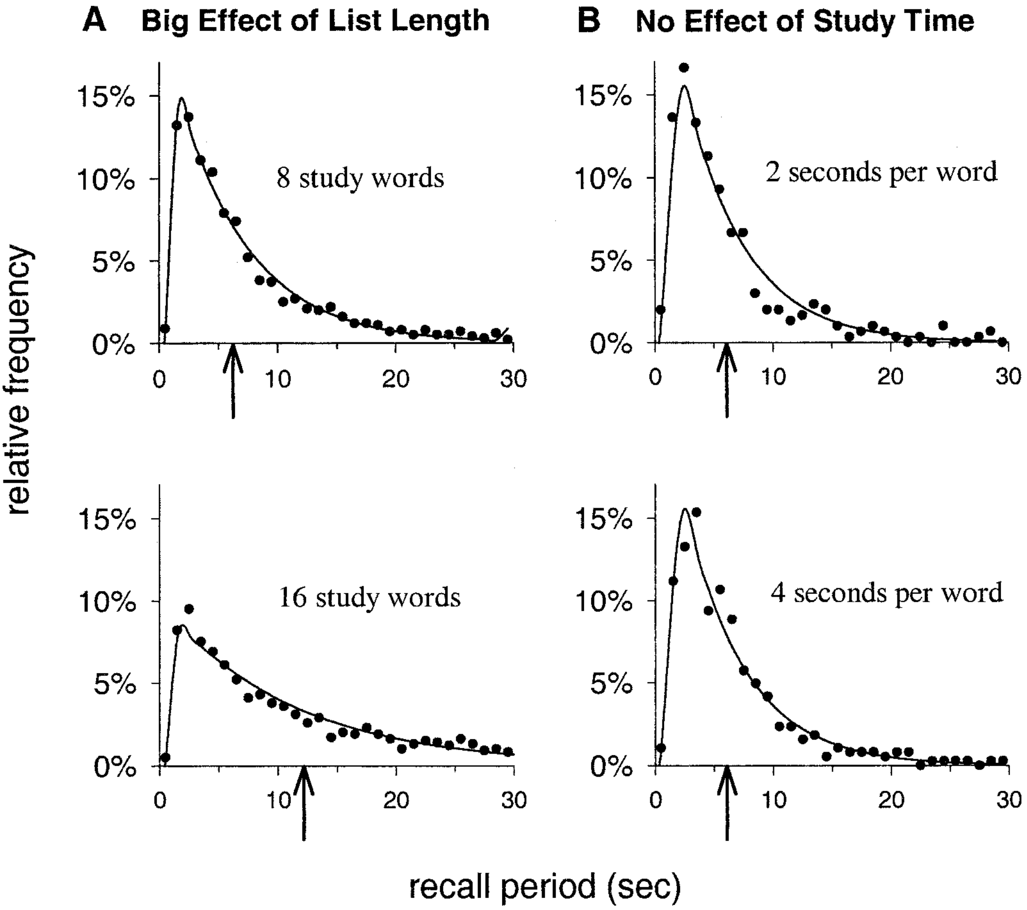

Figure 1. The effect of list length and study time on mean recall latency. (A) An increase in study list length increases mean recall latency, as indicated by the arrows (Rohrer, 1996). (B) An increase in study time has no effect on mean recall latency, as indicated by the arrows (Rohrer & Wixted, 1994).

average, researchers uniformly employ a different technique that effectively discounts the pause preceding the retrieval of the first response. Notably, this cannot be accomplished by simply subtracting the latency of the first response, because this latency includes the retrieval of the first response as well as the non-memorial processes. Therefore, researchers obtain parameter estimates of both the pause and the "true" mean recall latency by one of two equivalent curve-fitting techniques that yield an estimate of recall latency (or its reciprocal—the rate of approach to asympote; e.g., Bousfield & Sedgewick, 1944; Bousfield, Sedgewick, & Cohen, 1954; Gronlund & Shiffrin, 1986; Hermann & Murray, 1979; Indow & Togano, 1970; Maylor, Chater, & Brown, 2001; Payne, 1986; Roediger, Stellon, & Tulving, 1977; Roediger & Thorpe, 1978; Rohrer, 1996; Wixted, Ghadisha, & Vera, 1997; Wixted & Rohrer, 1994). The ex-Gaussian function is invariably the choice for recall latency distributions, and best-fitting ex-Gaussian curves are shown in Figure 1. An ex-Gaussian fit yields estimates of the pause and mean recall latency, and details are given in Appendix A. Estimates of the pause (μ) typically equal about 2 seconds, and estimates of mean recall latency (τ) vary dramatically as a function of the manipulation. Hence, this technique produces an estimate of mean recall latency that is approximately 2 seconds shorter than the observed mean recall latency. Notably, the distinction between the observed mean latency and the estimated mean latency is not critical when effect sizes are large, as in the present study.

Recall latency as a measure of search set size

The theoretical significance of recall latency is well illustrated by the manipulations of study list length and study time. As first shown by Roediger and Tulving (1979), longer study lists produce greater values of mean recall latency, and the data in Figure 1A are adapted from a replication of this effect (Rohrer, 1996). A longer study list also yields a greater recall *total*, of course, but an increase in recall total is not always associated with an increase in recall latency. For example, Rohrer and Wixted (1994) found that the manipulation of study time affects recall total (of course) without affecting recall latency, and the data in Figure 1B are adapted from this study.

The list length and study time manipulations are two of the numerous studies demonstrating that recall latency depends on the *study* total but not on the *recall* total, and this paradoxical result is consistent with the view that mean recall latency reflects search set size. That is, given that the search set generally includes all of the study items, the dependence of recall latency on study total ensures that recall latency also depends on search set size. Notably, this link between study list length and recall latency does not rule out the existence of extralist items within the search set—the focus of the present study. For example, if the search set included both study items and their strong semantic associates, study list length would remain proportional to search set size.

The construal of mean recall latency as a measure of search set size is further evidenced by the effect of proactive interference (PI) on recall latency. In the classic instantiation of this paradigm (e.g., Keppel & Underwood, 1962), participants study and attempt to recall three category members on each of several successive trials (e.g., three fruits, then three other fruits, and so on). Recall total declines across successive trials, because the rapid sequence of trials hinders the temporal discrimination of one study trial from another trial (cf. Nairne, Neath, Serra, & Byun, 1997). The study lists are effectively getting longer, as each list includes three additional items. This presumed increase in search set size across successive trials should therefore produce a concomitant increase in mean recall latency, and PI does, in fact, have this effect (Wixted & Rohrer, 1993).

In addition to these episodic memory studies, the relationship between mean recall latency and search set size also arises when subjects generate exemplars of a given category. In this semantic task, search set size is naturally assumed to depend on the number of known category exemplars. Hence, larger categories should produce greater values of mean recall latency, and this prediction has been verified several times (e.g., Bousfield & Sedgewick, 1944; Herrmann & Murray, 1979).

Further evidence is given by the effect of Alzheimer's disease (AD) on mean recall latency in the generation of category exemplars. AD is commonly believed to destroy the associations between representations within semantic memory, thereby reducing the number of exemplars within any given category (e.g., Heindel, Salmon, & Butters, 1990; Randolph, Braun, Goldberg, & Chase, 1993). Consequently, the construal of mean recall latency as a measure of search set size predicts that AD should reduce mean recall latency, and AD patients do, in fact, exhibit shorter mean recall latencies than demographically matched healthy controls (Rohrer, Wixted, Salmon, & Butters, 1995). By contrast, Huntington's disease is widely believed to leave semantic memory intact and instead retard the rate of processing (Heindel et al., 1990; Randolph et al., 1993), and, in accordance with this view, HD increases mean recall latency (Rohrer, Salmon, Wixted, & Paulsen, 1999). Notably, these last two studies demonstrate once again that the measures of recall latency and recall total may be correlated positively or negatively. The independence of these two measures is further illustrated in Table 1, which summarises the effects on these measures by the manipulations described earlier.

Summary

This study assesses whether category size affects mean recall latency during the episodic free recall

TABLE 1
The independence of recall latency and recall total

Manipulation	*Recall latency*	*Recall total*
Episodic		
Longer study list[1]	↑	↑
More study time[2]	—	↑
Build-up of proactive interference[3]	↑	↓
Semantic		
Greater category size[4]	↑	↑
Alzheimer's disease[5]	↓	↓
Huntington's disease[6]	↑	↓

[1] Roediger & Tulving, 1979
[2] Rohrer & Wixted, 1994
[3] Wixted & Rohrer, 1993
[4] Bousfield & Sedgewick, 1944
[5] Rohrer et al., 1995
[6] Rohrer et al., 1999

of categorically related study items. In particular, participants will attempt to freely recall a list of eight previously studied words drawn from the same category. In addition, participants will generate category exemplars from semantic memory, and, in this task, category size is known to positively predict mean recall latency. If recall latency increases with category size in the episodic task, the interpretation of mean recall latency as a measure of search set size suggests that episodic recall begins with a set of candidates that inappropriately includes semantic associates of the study items. If category size affects recall latency in the semantic task but not in the episodic task, the data would support the view that the breadth of search reflects the breadth of the episode rather than the breadth of the semantic category. Notably, the latter finding would also provide an empirical dissociation of episodic and semantic memory, regardless of the theoretical interpretation.

METHOD

Participants

A total of 36 undergraduates at the University of South Florida participated in return for course credit. This sample included 12 males and 24 females.

Materials

The episodic task and the semantic task utilised the same eight categories. The four small categories were South American Countries, Nuts, Wild Cats, and Furniture, and the four large categories were European Countries, Fruits, Birds, and Musical Instruments. The large categories were, in fact, normatively larger, as evidenced by the large effect of category size on recall total in the semantic task described in the Results section.

For the episodic task, the eight study words were drawn from the list of category exemplars listed in Appendix B. For the small-category condition, all eight of the listed exemplars were included. For the large-category condition, the eight study items were chosen randomly from the list of 24 category exemplars. This random selection was repeated for each participant, in order to ensure that the subset of items varied across participants.

Design

Participants completed two practice trials and eight scored trials, and each scored trial utilised one of the eight categories just listed. Participants alternated between episodic and semantic trials, and category size was counterbalanced. Each participant completed one episodic trial and one semantic trial with each of the four ‘‘kinds’’ of category (i.e., countries, edible plants, animals, and man-made objects). For example, if a participant was assigned ‘‘Fruits’’ in an episodic trial, that participant was necessarily assigned ‘‘Nuts’’ in a semantic trial.

Procedure

Participants were tested individually by personal computer in the presence of an experimenter. Each semantic trial began with the 3-s prompt, ‘‘Say examples of’’ followed immediately by the presentation of the category name. Participants then generated exemplars aloud during the 60-s recall period. Each episodic trial began with the 3-s prompt ‘‘Get ready to study’’ followed immediately by the eight study words at a rate of one per 3 s. The study phase was followed immediately by the appearance of 20 digit pairs at a rate of one per 1.5 s, and participants stated aloud the sum of each pair. A question mark then appeared on the screen, prompting participants to recall the study words in any order during the 60-s recall period.

After each response in either the semantic or episodic task, an experimenter immediately depressed a key on the computer keyboard. Although recall latency is sometimes measured with a voice-key apparatus during the free recall of monosyllabic items (Rohrer, 1996; Rohrer & Wixted, 1994), the key-press method is generally more reliable when measuring the continuous recall of polysyllabic items that arise during semantic memory retrieval. Although the key-press technique introduces a delay of about 200 ms that varies slightly across responses, this slight variability is meaningless once latencies are grouped into 1-s bins before further analysis.

RESULTS

Recall total

For both the episodic and semantic tasks, category size affected recall total in the usual manner. In

TABLE 2
Results

		Recall total		*Recall latency* (τ)		*Pause* (μ)	
Task	*Category size*	*mean*	*SE*	*mean*	*SE*	*mean*	*SE*
Episodic	Small	6.51	.13	8.32	0.42	1.90	.23
	Large	6.28	.17	8.80	0.45	2.07	.15
Semantic	Small	5.74	.30	12.90	0.76	1.95	.19
	Large	11.94	.57	22.64	1.12	2.00	.16
	Large (critical items)[1]	3.96	.94	22.16	1.88	1.98	.29

[1] This analysis included only the "critical" responses (see Discussion).

particular, category size did not affect recall total in the episodic task, $t(35) = 1.55$, n.s., yet the increase in category size significantly increased recall total in the semantic task, $t(35) = 13.07$, $p < .0001$. The mean recall totals and the standard errors for all four conditions are listed in Table 2.

Extralist intrusions and repetitions were extremely rare, as is typically observed in free recall studies. In the episodic task, the small and large category conditions included an average of 0.17 and 0.24 extralist intrusions per trial and 0.19 and 0.13 repetitions per trial, respectively. In the semantic task, the small and large category conditions included an average of 0.25 and 0.11 extracategory intrusions per trial and 0.03 and 0.10 repetitions per trial, respectively. The mean latency of either kind of error did not differ significantly as a function of category size in either the episodic or semantic task (all *t*s < 1).

Recall latency

Category size sharply affected mean recall latency in the semantic task but not in the episodic task, and this dissociation is clearly illustrated by the recall latency distributions in Figure 2. In the episodic task, for example, the fourth 1-s interval included about 12% of the responses, regardless of category size. In the semantic task, however, the fourth 1-s interval included more than 9% of the responses in the small category condition yet fewer than 6% of the responses in the large category condition.

This pattern of effects on mean recall latency was confirmed formally by comparing estimates of mean recall latency (τ) given by maximum likelihood estimation (MLE) fits of the ex-Gaussian function to the latency distributions. As noted in the Introduction and detailed in Appendix A, this technique provides an unbiased estimate of mean recall latency that excludes the brief pause preceding the retrieval of the first response. This analysis revealed that the increase in category size almost doubled mean recall latency in the semantic task without affecting mean recall latency in the episodic task, and the specific values are listed in Table 2. Statistical significance was assessed by a *t*-test comparison of parameter estimates, and this *t*-test incorporates the asymptotic standard errors provided by the Hessian matrix of second partial derivatives (cf. Maindonald, 1984; Ratkowski, 1983). This test revealed that the difference in τ was statistically significant in the semantic task, $t(114) = 8.67$, $p < .0001$, but not in the episodic task, $t(114) < 1$, n.s.

DISCUSSION

The time course of continuous recall was shown to be dependent on category size for the semantic memory task but not for the episodic memory task. Specifically, the generation of exemplars from a semantic category produced mean latencies that increased as a function of category size, whereas the episodic recall of eight previously studied, categorically related words produced mean latencies that were unaffected by the natural size of the category (Figure 2). Thus, although both the episodic and semantic tasks required the continuous recall of categorically related items, the time course of recall was affected differentially by the manipulation of category size. This finding provides a dissociation of episodic and semantic memory.

This dissociation is probably best described as functional rather than structural. For instance, episodic memory retrieval may rely on the representations within semantic memory, and the ability to constrain search to intralist items may simply reflect the capability to exclude repre-

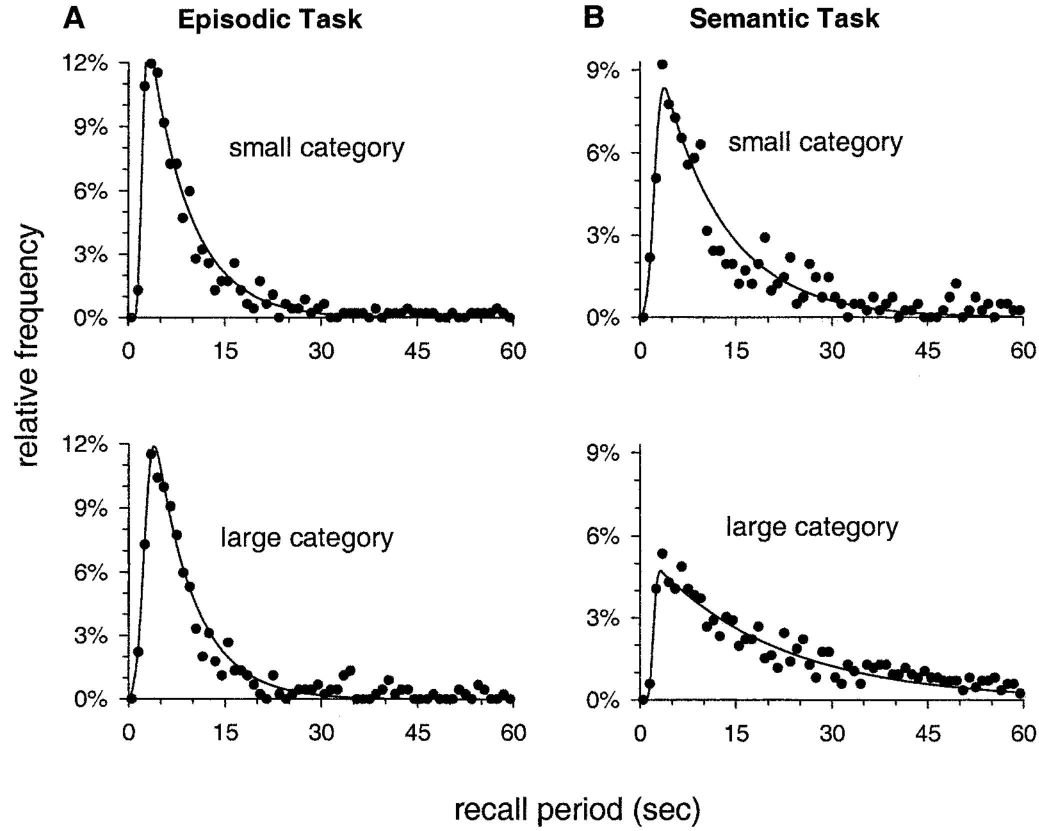

Figure 2. Observed recall latency distributions. (A) In the episodic task, category size does not affect mean recall latency (Table 2). (B) In the semantic task, greater category size sharply increases mean recall latency (Table 2).

sentations that have not been recently activated. For example, when recalling eight recently studied kinds of fruit from among the larger set of all kinds of fruit, the search set may include only the eight recently activated "warm" representations. Hence, these data allow for the possibility that one or more structures mediate both kinds of memory systems.

This dissociation is given added theoretical significance when mean recall latency is construed as a measure of search set size. As described fully in the Introduction, this account presumes that search begins once a limited number of possibilities are delimited, with larger search sets producing longer mean recall latencies. In episodic free recall, the search set includes the study items, as consistent with the finding that mean recall latency varies directly with the number of study items. Likewise, the generation of semantic category exemplars is believed to involve a search through the semantic representations of every exemplar within the category, as consistent with the finding that these semantic tasks yield mean recall latencies that vary with category size.

By this interpretation of mean recall latency, the null effect of category size on latency in the episodic task suggests that episodic free recall relies on a search that excludes extralist category exemplars. That is, because the manipulation of category size is necessarily a manipulation of the number of *unstudied* category exemplars, the null effect of category size suggests that these extralist items were excluded from the search set in the episodic task. By this view, the free recall of eight previously studied "fruits" relies on a search that efficiently excludes the many extralist "fruits", despite the semantic associations between these extralist items and the intralist items. In summary, then, this view provides a parsimonious account of the observed dissociation of episodic and semantic continuous recall. The breadth of semantic search depends on the breadth of the semantic category, and the breadth of episodic search depends on the breadth of the episode.

The possibility that episodic retrieval is limited to a search that regularly excludes extralist semantic associates is also consistent with the salient rarity of extralist errors in free recall. In the free recall data reported here, for instance, participants produced an average of less than one extralist intrusion in every *four* trials, regardless of category size. Yet the likelihood of extralist intrusions can be increased dramatically by certain manipulations. For example, in the well-known false recall procedure popularised by Roediger and McDermott (1995), study lists that contain many strong associates of the same extralist word often lead to the subsequent false recall of the critical extralist word. This intriguing and oft-replicated finding is at odds with the results reported here, as participants in the present study rarely recalled extralist items despite the semantic associations between the intralist and extralist category members. Perhaps the two findings can be reconciled by noting that, although episodic memory retrieval is generally constrained to the study items, the delimiting process can be overwhelmed when a single extralist item is a strong associate of every intralist item.

The relative strength model

This search set interpretation of the data reported here and other recall latency data is consistent with a relative strength model of retrieval. This model attributes an item's recall latency to a competition between items, with each item competing against all of its peers. The earliest versions of this model required items of equal strength, (e.g., McGill, 1963; Shiffrin, 1970), but the relaxation of this untenable assumption has since been shown to improve the quality of fit (cf. Rohrer, 1996; Vorberg & Ulrich, 1987; Wixted et al., 1997).

The relative strength model attributes an item's recall *likelihood* to its absolute strength and its *latency* to its relative strength, and the distinction is illustrated by the hypothetical data in Figure 3. In Figure 3A, the manipulation of list length has no effect on the absolute strength of any item, because the manipulation does not affect the amount of study time allotted to each item. Yet the increase in the number of study words decreases each item's relative strength, because each item's share of the total strength depends on the total number of study items. In essence, longer study lists produce bigger search sets, and larger search sets require greater average search times. By contrast, the hypothetical data in Figure 3B illustrate that an increase in study time increases each item's absolute strength without affecting its relative strength. The increase in absolute strength increases recall total, because a greater number of items will exceed the threshold needed for recovery. However, because *every* item is stronger, each item's share of the total strength remains constant. Hence, each item's relative strength remains constant, and mean recall latency is therefore unaffected. In effect, greater study time increases the strength of all items in the search set, but the average time to find each item remains unchanged.

Ironically, then, the relative strength model predicts that a weak item within a search set of a few other equally weak items will be recalled more quickly than a strong item within a search set of many equally strong items. Hence, short, quickly presented study lists should yield shorter mean recall latencies than long, slowly presented study lists, and this prediction has been confirmed as well (Rohrer & Wixted, 1994).

Recall latency and recall total

The present study revealed that category size affected recall latency and recall total in the same manner, and this naturally raises the question of whether the effects on latency were simply artifacts of the effects on recall total. That is, if recall latency intrinsically depended on the number of items recalled, the observed effects on recall latency would be dictated by the effects on recall total. However, this rival hypothesis is countered by several arguments.

First, the studies described in the Introduction demonstrate that the measures of recall total and recall latency may be correlated positively, negatively, or not at all. For episodic free recall, for instance, the two measures correlate positively when list length varies and correlate negatively with the build-up of PI. With semantic memory, the two measures correlate positively when AD patients are compared to healthy controls, whereas the progression of HD causes a decline in recall total that is accompanied by an increase in recall latency.

In addition, the data rule out the possibility that the two measures were correlated positively because of limitations in the rate of speech. That is, the two measures are necessarily correlated when subjects recall words as quickly as they can pronounce the words, and this occurs in the recall from rote memory. For example, the recall of the 12 months of the year *does* produce a positive

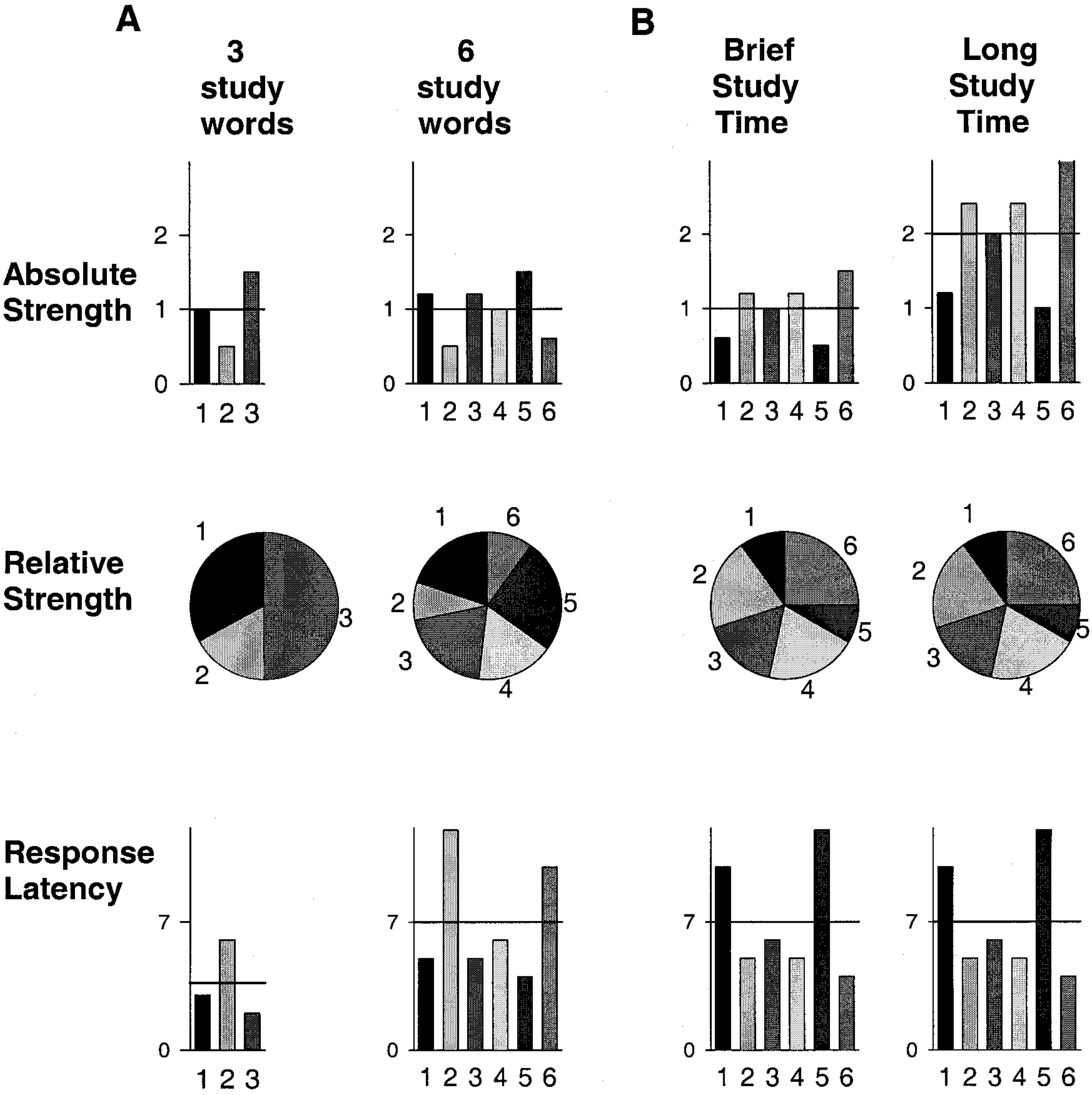

Figure 3. Hypothetical illustration of relative strength model. (A) List length manipulation. Absolute strength is unaffected as study time remains constant, but the increase in study item total reduces relative strength and consequently slows mean recall latency. (B) Study time manipulation. Absolute strength is increased with more study time, but relative strength and mean recall latency remain unaffected because the number of study items remains constant. Each horizontal line represents the mean of the values given in that plot.

correlation between recall total and recall latency, because the latencies of all but the first response are delayed by the need to first verbalise the earlier responses. Yet the present study required a "search" of memory, and, consequently, the rate of recall was much slower than the optimal rate of pronunciation. In the condition with the greatest recall total, for instance, participants produced an average of only 12 words with a mean recall latency of about 22 s. If these 12 words had been recalled at the rate of pronunciation, however, mean recall latency would have equalled less than 7 s (assuming a conservative rate of one word per second). Hence, the longer latencies of the items recalled late in the recall period reflect the time needed to retrieve these items rather than a bottleneck in verbal output.

Finally, the independence of mean recall latency and recall total is further illustrated by a post-hoc analysis of the large-category semantic data. In particular, although an increase in category size increased both recall latency and recall total in the semantic task, the increase in recall latency holds regardless of whether a large proportion of recall items are excluded. In particular, this post-hoc analysis included only a randomly chosen one-third of the responses in the large-category semantic task, just as the large-category episodic study lists included a randomly chosen one-third of the category exemplars. Naturally, this exclusion of items reduced recall total to one third of its original amount, and this reduced total was significantly less than the recall total for the small-category semantic task (see Table 2). Yet mean recall latency for these "critical" responses in the large-category semantic task was virtually unchanged by the exclusion of the "non-critical" responses (see Table 2). This result is counter to the natural intuition, perhaps, but it might be clarified by an analogy. If a distribution of test scores for a large class has a mean score of, say, 80, the mean score should not be affected by the removal of a randomly selected subset of scores.

The effect of semantic associations on episodic memory

Notably, the finding reported here is not inconsistent with previous findings revealing an effect of semantic associations on episodic free recall. For example, Nelson, McEvoy, and Schreiber (1990) report that free recall total can be affected dramatically by the number of words that are strongly associated to the study items, but this finding is orthogonal to the outcome of the present study (D.L. Nelson, personal communication, 12 June 2001).

Nor do the present data conflict with the robust phenomenon of clustering in free recall that clearly reflects the role of semantic associations. In two such studies, for instance, a 25-word study list included five members from each of five categories in a random order, yet recall order was clustered by category: several fruits, then several vegetables, and so on (Patterson et al., 1971; Pollio et al., 1969). Yet these clusters may simply reflect the use of a different search set for each category. In both of these studies, in fact, the rate of recall was fastest at the beginning of each cluster, just as rates of recall are faster at the beginning of the recall period after the study of words from only one category.

In conclusion, the results of these studies and the study reported here are consistent with a view that reveals memory retrieval to be remarkably efficient. By this account, episodic retrieval relies on a search set that excludes the extraneous semantic associates of the episodically related items while simultaneously exploiting the semantic associations between these items.

Manuscript received 4 July 2001
Manuscript accepted 24 October 2001

REFERENCES

Bousfield, W.A., & Sedgewick, C.H. (1944). An analysis of restricted associative responses. *The Journal of General Psychology*, *30*, 149–165.

Bousfield, W.A., Sedgewick, C.H., & Cohen, B.W. (1954). Certain temporal characteristics of the recall of verbal associates. *American Journal of Psychology*, *67*, 111–118.

Gronlund, S.D., & Shiffrin, R.M. (1986). Retrieval strategies in recall of natural categories and categorized lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 550–561.

Heindel, W.C., Salmon, D.P., & Butters, N. (1990). Pictorial priming and cued recall in Alzheimer's and Huntington's disease. *Brain and Cognition*, *13*, 282–295.

Herrmann, D.J., & Murray, D.J. (1979). The role of category size in continuous recall from semantic memory. *Journal of General Psychology*, *101*, 205–218.

Indow, T., & Togano, K. (1970). On retrieving sequency from long-term memory. *Psychological Review*, *77*, 317–331.

Keppel, G., & Underwood, B.J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior*, *1*, 153–161.

Maindonald, J.H. (1984). *Statistical computation*. New York: Wiley.

Maylor, E.A., Chater, N., & Brown, G.D.A. (2001). Scale invariance in the retrieval of retrospective and prospective memories. *Psychonomic Bulletin & Review*, *8*, 162–167.

McGill, W.J. (1963). Stochastic latency mechanisms. In R.D. Luce, R.R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1, pp. 309–360). New York: Wiley.

Nairne, J.S., Neath, I., Serra, M., & Byun, E. (1997). Positional distinctiveness and the ratio rule in free recall. *Journal of Memory and Language*, *37*, 155–166.

Nelson, D.L., McEvoy, C.L., & Schreiber, T.A. (1990). Encoding context and retrieval conditions as determinants of the effects of natural category size. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 31–41.

Patterson, K.E., Meltzer, R.H., & Mandler, G. (1971). Inter-response times in categorized free recall. *Journal of Verbal Learning and Verbal Behavior*, *10*, 417–426.

Payne, D.G. (1986). Hypermnesia for pictures and words: Testing the recall level hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 16–29.

Pollio, H.R., Richards, S., & Lucas, R. (1969). Temporal properties of category recall. *Journal of Verbal Learning and Verbal Behavior*, *8*, 529–536.

Randolph, C., Braun, A.R., Goldberg, T.E., & Chase, T.N. (1993). Semantic fluency in Alzheimer's, Parkinson's, and Huntington's disease: Dissociation of storage and retrieval failures. *Neuropsychology*, *7*, 82–88.

Ratkowski, D.A. (1983). *Nonlinear regression modeling: A unified practical approach*. New York: M. Dekker.

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., Stellon, C.C., & Tulving, E. (1977). Inhibition from part-list cues and rate of recall. *Journal of Experimental Psychology: Human Learning and Memory*, *3*, 164–188.

Roediger, H.L., & Thorpe, L.A. (1978). The role of recall time in producing hypermnesia. *Memory & Cognition*, *6*, 296–305.

Roediger, H.L., & Tulving, E. (1979). Exclusion of learned material from recall as a post-retrieval operation. *Journal of Verbal Learning and Behavior*, *18*, 601–615.

Rohrer, D. (1996). On the relative and absolute strength of a memory trace. *Memory & Cognition*, *24*, 188–202.

Rohrer, D., Salmon, D.P., Wixted, J.T., & Paulsen, J.S. (1999). The disparate effects of Alzheimer's disease and Huntington's disease on semantic memory. *Neuropsychology*, *13*, 381–388.

Rohrer, D., & Wixted, J.T. (1994). An analysis of latency and interresponse time in free recall. *Memory & Cognition*, *22*, 511–524.

Rohrer, D., Wixted, J.T., Salmon, D.P., & Butters, N. (1995). Retrieval from semantic memory and its implications for Alzheimer's disease. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1127–1139.

Shiffrin, R.M. (1970). Memory search. In D.A. Norman (Ed.), *Models of human memory* (pp. 375–447). New York: Academic Press.

Tulving, E. (1986). What kind of a hypothesis is the distinction between episodic and semantic memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 307–311.

Tulving, E. (1993). What is episodic memory? *Current Directions in Psychological Science*, *2*, 67–70.

Vorberg, D., & Ulrich, R. (1987). Random search with unequal search rates: Serial and parallel generalizations of McGill's model. *Journal of Mathematical Psychology*, *31*, 1–23.

Wixted, J.T., Ghadisha, H., & Vera, R. (1997). Recall latency following pure- and mixed-strength lists: A direct test of the relative strength model of free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 523–538.

Wixted, J.T., & Rohrer, D. (1993). Proactive interference and the dynamics of free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 1024–1039.

Wixted, J.T., & Rohrer, D. (1994). Analyzing the dynamics of free recall: An integrative review of the empirical literature. *Psychonomic Bulletin & Review*, *1*, 89–106.

APPENDIX A

The ex-Gaussian distribution is given formally by

$$\mathrm{f(x)} = \left(\frac{e^{\frac{\mu}{\tau}-\frac{x}{\tau}+\frac{\sigma^2}{2\tau^2}}}{\tau\sqrt{2\pi}}\right)\int_{-\infty}^{\left(\frac{x}{\sigma}-\frac{\mu}{\sigma}-\frac{\sigma}{\tau}\right)} e^{-\frac{y^2}{2}}dy\ . \qquad \text{(A1)}$$

The ex-Gaussian distribution is a convolution of a normal distribution and an exponential distribution, and a derivation is given in Rohrer and Wixted (1994). A latency distribution is ex-Gaussian if the latency comprises a normally distributed stage and an exponentially distributed stage. In the case of recall latency distributions, the normal component describes the initial pause, and it is characterised by its mean (μ) and its standard deviation (σ). The exponential component describes the retrieval phase and is defined solely by its mean, τ. These parameter estimates are unbiased when the best fitting function is found by maximum likelihood estimation, and this technique produced the estimates given in Table 2.

APPENDIX B

South American countries	**Nuts**	**Wild cats**	**Furniture**
Argentina	almond	bobcat	bed
Bolivia	cashew	cheetah	bookcase
Chile	chestnut	jaguar	chair
Colombia	filbert	leopard	desk
Ecuador	macadamia	lynx	dresser
Paraguay	pecan	panther	sofa
Uruguay	pistachio	puma	stool
Venezuela	walnut	tiger	table

European countries	**Fruits**	**Birds**	**Musical instruments**
Albania	avocado	crow	bagpipes
Austria	apricot	dove	banjo
Belgium	apple	duck	bassoon
Bulgaria	banana	eagle	bongo
Croatia	blueberry	falcon	bugle
Denmark	cantaloupe	flamingo	cello
Estonia	cherry	goose	clarinet
Finland	coconut	hawk	drum
France	fig	jay	flute
Germany	kiwi	lark	guitar
Greece	kumquat	oriole	harmonica
Hungary	lemon	ostrich	harp
Ireland	lime	parrot	kazoo
Italy	mango	pelican	oboe
Latvia	melon	penguin	piccolo
Norway	nectarine	pheasant	saxophone
Poland	orange	pigeon	sitar
Portugal	peach	raven	triangle
Romania	pear	robin	trombone
Serbia	pineapple	sparrow	trumpet
Spain	plum	stork	tuba
Sweden	prune	swan	ukulele
Switzerland	tangerine	turkey	violin
Yugoslavia	watermelon	vulture	viola