Journal of Experimental Psychology:
Learning, Memory, and Cognition
1995, Vol. 21, No. 5, 1127–1139

0278-7393/95/$3.00

Retrieval From Semantic Memory and Its Implications for Alzheimer's Disease

Doug Rohrer, John T. Wixted, David P. Salmon, and Nelson Butters
University of California, San Diego

In 3 experiments, participants generated category exemplars (e.g., kinds of fruits) while a voice key and computer recorded each response latency relative to the onset of responding. In Experiment 1, mean response latency was faster when participants generated exemplars from smaller categories, suggesting that smaller mental search sets result in faster mean latencies. In Experiment 2, a concurrent secondary task increased mean response latency, suggesting that slowed mental processing results in slower mean latencies. In Experiment 3, the mean response latency of Alzheimer's participants was faster than that of elderly controls, which is consistent with the idea that the semantic memory impairments of Alzheimer's disease patients stem primarily from a reduction in available items (as in Experiment 1) rather than retrieval slowing (as in Experiment 2).

Although Alzheimer's disease is most often associated with episodic memory impairment, the semantic memory deficits of Alzheimer's disease (AD) patients are equally striking, even in the earliest stages of the disease. Specifically, AD patients exhibit impaired verbal production, whereas their verbal comprehension remains relatively intact until latter stages of the disease (cf. Martin & Fedio, 1983). Thus, semantic information is, in some sense, still intact yet unable to be accessed. The cause of this word-finding difficulty is a focus of extensive research and debate. The *structure-loss* hypothesis attributes this inaccessibility to the loss of interitem associations between items within semantic memory, whereas the *retrieval-slowing* hypothesis posits that semantic memory structures are intact but often inaccessible.

There is a conspicuous difference between the data supporting these two hypotheses. The evidence for structure loss relies heavily on the number of category exemplars (e.g., kinds of fruits) produced by AD and normal control (NC) participants during a given period of time (usually 1 min). None of the evidence for the structure-loss hypothesis includes response time data. Conversely, the evidence for retrieval slowing is composed entirely of response times, none of which were measured during a category fluency task. Bridging this gap, the analyses presented in this article focus on retrieval latency in a semantic fluency task. Before considering these results, the structure-loss hypothesis, the retrieval-slowing hypothesis, and the nature of retrieval latency in healthy individuals are first described in more detail.

Doug Rohrer and John T. Wixted, Department of Psychology, University of California, San Diego; David P. Salmon, Department of Neurosciences, University of California, San Diego; and Nelson Butters, Department of Psychiatry, University of California, San Diego.

This research was supported in part by National Institute of Mental Health (NIMH) Training Grant MH-14268, National Institute on Aging Grant AG 05131, NIMH Grant MH48819, and the Medical Research Service of the Department of Veterans Affairs.

We thank Mari Chernow, Eric Hanna, Nicole Hansen, and Robie Yamamoto for their assistance in the laboratory; Hal Pashler and Ling-Po Shiu for their advice on Experiment 2; and Scott Gronlund for his comments.

Correspondence concerning this article should be addressed to Doug Rohrer, who is now at the Department of Psychology, George Washington University, 2125 G Street, NW, Washington, DC 20052.

Structure-Loss Hypothesis

Many researchers believe that the word-finding impairments that are observed in AD patients reflect a breakdown in the associative networks that characterize semantic memory. Thus, an AD patient who is able to recognize a particular word may not be able to retrieve it without the interitem associations between that word and others. Consider, for example, the oft cited hierarchy of "vertical" associations shown in Figure 1. A broad category such as *animals* subsumes *smaller* categories such as *farm animals,* which in turn subsume exemplars such as *cow* and *pig.* ("Horizontal" associations such as *cow–pig* exist as well.) That these vast networks of interitem associations facilitate retrieval is supported by the phenomenon of clustering. Specifically, as participants produce exemplars of a relatively broad category (e.g., *animals*), items from the same subset (e.g., *farm animals*) are often retrieved consecutively and with brief interresponse times (cf. Bousfield & Sedgewick, 1944; Graesser & Mandler, 1978; Pollio, 1964).

The primary evidence for the loss of semantic structure derives from the manner in which AD participants produce exemplars belonging to a given category. First, for example, Martin and Fedio (1983) and Tröster, Salmon, McCullough, and Butters (1989) had AD participants generate objects found in a supermarket, a category which subsumes numerous subcategories (e.g., fruits, vegetables, and so forth), for a period of 60 s. AD participants in both studies produced only about two items per subcategory, significantly fewer than NC participants, even though AD participants switched to new subcategories less often than the controls. Second, both groups of researchers reported that AD participants were more likely to produce superordinate responses (e.g., "vegetables") rather than specific items (e.g., "corn"), suggesting that the loss of interitem associations is "bottom-up." Neither finding is readily explained by a retrieval deficit.

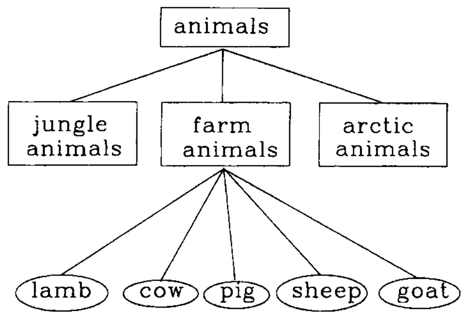

Figure 1. Hierarchical network model of semantic memory.

In three similar studies, researchers measured not only the production of exemplars belonging to a semantic category but also the production of words beginning with a specified letter (Butters, Granholm, Salmon, Grant, & Wolfe, 1987; Mickanin, Grossman, Onishi, Auriacombe, & Clark, 1994; Monsch et al., 1994). In each study, AD participants exhibited greater impairment in the category fluency than in the letter fluency task, relative to controls. These results are consistent with the hypothesized loss of interitem associations, given the reasonable assumption that category fluency, unlike letter fluency, greatly depends on interitem associations (both hierarchical and horizontal).

In another study of category fluency, Randolph, Braun, Goldberg, and Chase (1993) presented participants with names of *large* categories (e.g., *animals*) with or without four subcategory cues (e.g., *pets, jungle animals, water animals,* and *farm animals*). Whereas the presence of the retrieval cues significantly increased the number of items produced by Huntington's disease (HD) and Parkinson's disease participants, these subcategory names had no effect on the response total of AD participants. If the hierarchical associations between exemplars and their corresponding subcategory names were intact in these AD participants, the presentation of these subcategory names as retrieval cues should have facilitated retrieval of the appropriate exemplars.

Finally, Chertkow and Bub (1990) showed that the category exemplars that cannot be retrieved by AD participants in a category fluency task are generally the same items that cannot be accessed by a more direct paradigm. These authors probed the extent to which each participant understood the meaning of each of 130 words by asking up to a dozen questions about each word and its attributes. On the basis of this questioning, the researchers divided the 130 words into two groups for each participant: intact and degraded. When these participants were later asked to generate category exemplars (all of which had been probed earlier), almost 90% of the intact items, but only 12% of the degraded items, were produced. Thus, the items that could not be accessed by probing were roughly the same items that were not retrieved in the category fluency task. That the same items could not be accessed by either method is consistent with a loss of structure. Presumably, impairments in retrieval would produce deficits that vary across different methodologies.

Retrieval-Slowing Hypothesis

Whereas the structure-loss hypothesis may be construed as a storage deficit, Nebes and his associates argued that semantic impairment in Alzheimer's disease reflects a retrieval deficit. As described by Nebes, Alzheimer's disease is characterized by a "preserved" semantic structure (Nebes & Brady, 1990, p. 574) and a "generalized slowing of cognitive processing" (Nebes & Brady, 1992, p. 317).

As evidence for these views, Nebes and his associates presented response time data for a wide array of semantic retrieval tasks. In one typical experiment, participants heard a sentence without the final word and were asked to complete the sentence with a word that "made sense" (Nebes, Boller, & Holland, 1986). The difficulty of the sentence completions varied across three levels. The easier sentences had fewer acceptable final words ("Most cats see well at ____") than the more difficult, less restrictive sentences ("In the distance they heard a ____"). At each level of difficulty, the mean response time for AD participants was significantly slower than that of controls, which is consistent with the retrieval-slowing hypothesis. In addition, mean response-time increased as a function of sentence-completion difficulty to the same degree for both groups of participants. Nebes interpreted this increase in response time for AD participants as evidence for the preservation of associative networks within semantic memory. Specifically, Nebes argued that the highly restrictive nature of the easier sentence completions results in the activation of fewer words within semantic memory, thereby facilitating faster response times for participants with preserved semantic associations. Thus, if the associative networks were not intact in AD participants, their response times should not be affected by the restrictiveness of the task. Nebes and his associates used this same reasoning in their analysis of each of their experiments: word naming with semantic priming (Nebes, Martin, & Horn, 1984), sentence completion with semantic priming and yes–no judgments of category–exemplar pairs (Nebes et al., 1986), word–nonword judgments with semantic priming and word naming with semantic priming (Nebes, Brady, & Huff, 1989), yes–no judgments about related–nonrelated word pairs (Nebes & Brady, 1990), and yes–no judgments about sensible–nonsensible sentence completions (Nebes & Brady, 1991). As described above, the response times of AD participants in each of these experiments, though slower than the response times of controls, were affected by the degree of difficulty in roughly the same manner as the response times of controls.

Notably, Nebes did not explain how a retrieval deficit might account for the markedly few exemplars generated by AD participants in a category fluency task. If AD participants were simply slower, then response totals should not be affected. However, it could reasonably be argued that this slowing might reach a critical threshold beyond which certain items can no longer be retrieved. Furthermore, in all of the category fluency experiments cited above researchers used categories that were much too large, given their use of recall periods lasting only 60 s. For example, there are hundreds of supermarket items and thousands of words beginning with the letter *s*. Thus, the use of longer recall periods, or better yet, smaller categories, might have eliminated much of the difference in response totals between AD and NC participants.

Time Course of Retrieval

Over the last 50 years, sporadic attention has been given to response latency in tasks requiring the retrieval of category exemplars from semantic memory (Baddeley, Lewis, Eldridge, & Thomson, 1984; Bousfield & Sedgewick, 1944; Gruenewald & Lockhead, 1980; Herrmann & Chaffin, 1976; Herrmann & Murray, 1979; Indow & Togano, 1970; Johnson, Johnson, & Mark, 1951; Metlay, Handley, & Kaplan, 1971; Wixted & Rohrer, 1994) and response latency in the free recall from episodic memory (Bousfield, Sedgewick, & Cohen, 1954; Gronlund & Shiffrin, 1986; Payne, 1986; Roediger & Payne, 1985; Roediger, Payne, Gillespie, & Lean, 1982; Roediger, Stellon, & Tulving, 1977; Roediger & Thorpe, 1978; Roediger & Tulving, 1979; Rohrer & Wixted, 1994; Wixted & Rohrer, 1993). In most of these analyses, a dependent measure that derives from the mathematical form of the decline in retrieval as a function of time was used.

As an illustration of this decline, a hypothetical plot of retrieval as a function of time is given in Figure 2. Each data point represents the number of correct items produced in each 5-s interval of a 60-s recall period. Thus, the filled-circle participants produced more responses between 10 and 15 s than the open-circle participants but fewer items between 40 and 45 s. This decline in output is, in practice, well described by a simple exponential equation,

$$r(t) = (N/\tau)^{*}e^{-t/\tau},$$

where $r(t)$ represents the number of items recalled at time t. N, representing asymptotic recall (the estimated number of items that could be produced given unlimited time), and τ, representing the mean latency of those N items, are parameter estimates derived from fitting the equation to the data.

Asymptotic recall (N) equals observed total recall if participants are given enough time to produce all of their available exemplars. As seen in Figure 2, participants are still producing items at the end of the recall period, especially those in the open-circle condition. The estimation of asymptotic recall can

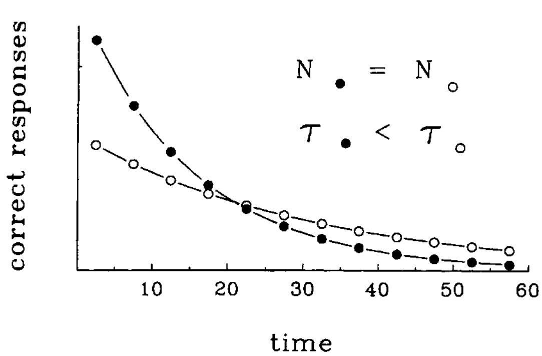

Figure 2. Idealized data showing the exponential decline in mean items produced in each 5-s bin. N represents asymptotic recall, and τ represents mean latency.

therefore provide a more accurate measure of the number of available items, especially when participants are able to make significant progress beyond the end of the recall period. In the aforementioned letter fluency task, for example, participants can essentially produce words beginning with a specified letter for at least 5 min.

The estimate of mean latency (τ) provides a measure of the most salient (and most ignored) feature of retrieval from semantic memory: the remarkably long response latencies. Specifically, latency for a recalled item is defined as the time elapsed since the onset of the recall period (not the time since the previous response). Mean latency is simply the average of the response latencies. For example, if a participant produced three category exemplars—say, at 5 s, 10 s, and 30 s—then the mean latency for that trial equals 15 s. As an illustration of mean latency, consider again the idealized data shown in Figure 2. Mean latency for the filled-circle participants, who produce a high proportion of their responses early on, is less than that of the open-circle participants. A second way to estimate mean latency from such plots is to compare the "half-life" of each condition, the time needed for output to decline by 50%. In Figure 2, the height of the filled-circle function declines by 50% in about 10 s, whereas the height of the open-circle function declines by 50% in about 20 s. Of importance, if decline is exponential, the estimate of mean latency (τ) equals the "observed" mean latency only if participants complete responding before the end of the recall period. If not, comparison of observed mean latency across conditions can be misleading.

Theoretically, exponentially declining rates of retrieval are produced by the well-known random-search model, as first noted by McGill (1963). Though the random-search model can parsimoniously be envisioned as a parallel process (cf. Rohrer & Wixted, 1994), it is typically presented as a serial process. According to the serial interpretation, the retrieval cue (e.g., *farm animals*) delimits a mental search set that contains the relevant items (e.g., *lamb, sheep, cow,* and so forth). Exemplars are randomly sampled one at a time, at a constant rate. Each item has the same probability of being sampled, and this probability holds constant throughout the recall period. Each sampled exemplar is immediately recognized as either a not-yet-sampled exemplar (and then retrieved into consciousness) or a previously sampled exemplar (and then ignored). As the number of not-yet-sampled items decreases, the number of items retrieved in each, say, 5-s bin correspondingly declines throughout the recall period.

Numerous elaborations of this rather simplified model of retrieval have been proposed. These modifications are designed to address retrieval phenomena other than recall latency, the focus of our investigation. For example, clustering can be explained by allowing for a hierarchical search set (Graesser & Mandler, 1978; Gruenewald & Lockhead, 1980; Herrmann & Pearle, 1981; Pollio, 1964). In this manner, a search set (e.g., *animals*) may include a sub-search set (e.g., *mammals*) that may itself include exemplars that are closely related (e.g., *dolphin–whale*). In fact, Herrmann and Pearle showed mathematically that such a sampling hierarchy, depending on exact assumptions, can still produce perfect exponential decline in retrieval output. Similar mathematical modifications have incorporated variations in item strength (Vorberg & Ulrich, 1987) and individual differences (Morrison, 1979). The general consensus among these authors, however, is that these mathematical extensions should be considered when one of these factors is the focus of study or is likely to change substantially across experimental conditions.

According to the most straightforward model of random sampling, mean latency depends on both the breadth of search and speed of processing, just as intuition would suggest. More formally,

$$\tau = St^*,$$

where S equals the number of items within the search set, and t^* equals the time needed to sample a single item from the search set (cf. McGill, 1963; Rohrer & Wixted, 1994). Therefore, the average time needed to retrieve the items within the search set increases when either the size of the search set increases or the duration of each random sampling increases.

The measure of mean latency (τ) serves as a direct test of the structure-loss and retrieval-slowing hypotheses. As illustrated in Figure 3, the loss of hierarchical associations would result in a smaller search set so that the mean latencies for AD participants would be faster than that of NC participants. In contrast, slowed processing would slow sampling times so that the mean latencies for AD participants would be slower than those of NC participants, which is perhaps the natural intuition. Thus, the two hypotheses actually predict opposite effects on the same dependent measure, mean latency. As an illustration of these predicted findings, Figure 4 shows idealized latency data for the structure-loss hypothesis (A) and the retrieval-slowing hypothesis (B). In both Figures 4A and 4B, asymptotic recall (N) for AD participants is one third that of NC participants (as typically found). However, in Figure 4A, τ_{AD} equals one half τ_{NC}, whereas in Figure 4B, τ_{AD} is twice τ_{NC}. Thus, Figure 4A depicts a faster mean latency for AD participants as would be expected, given the loss of interitem associations between the category retrieval cue and many of its exemplars (i.e., a smaller search set). Figure 4B depicts a longer mean latency for AD participants as would be expected if their retrieval was significantly slowed. As discussed earlier, these differences in mean latency can visually be estimated by comparing half-lives. In Figure 4A, output by AD participants declines by one half in just over 5 s, whereas output by NC participants declines by one half in about 15 s. In Figure 4B, however, the decline by AD participants requires much more time than controls.

Though the dependence of mean latency on search-set size and processing speed is supported by both intuition and the random-search model, Experiments 1 and 2 provided empirical evidence. In Experiment 1, college age participants produced exemplars from either small or large categories, a manipulation that should affect the number of items within the search set (Indow & Togano, 1970). In Experiment 2, college age participants produced exemplars with or without having to perform a concurrent secondary task, a manipulation that should affect processing speed. In Experiment 3, AD and NC participants produced category exemplars, and the results are compared with those predicted by the structure-loss and retrieval-slowing hypotheses, as described above.

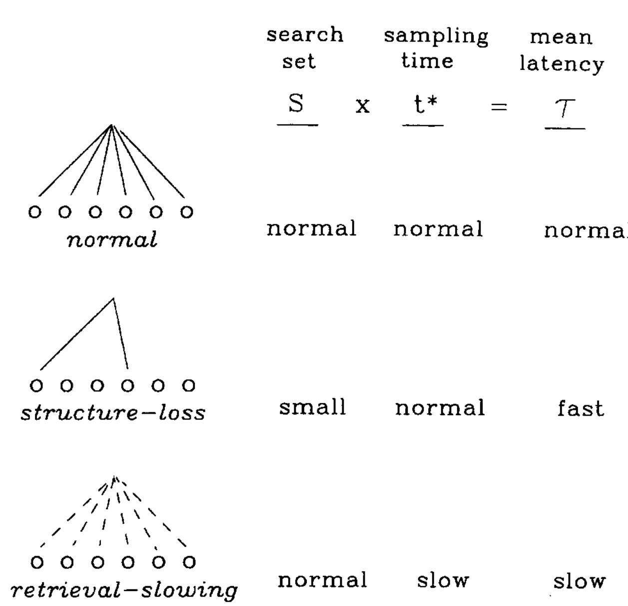

Figure 3. The effects of structure loss and retrieval slowing on mean latency.

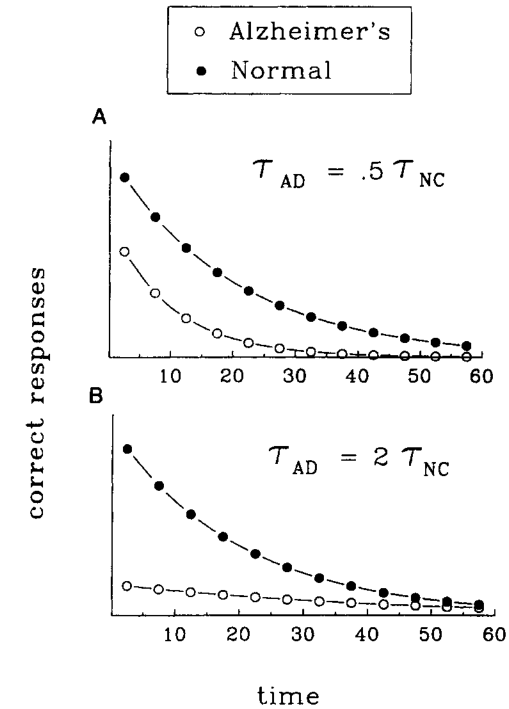

Figure 4. Idealized recall output for each 5-s bin, declining exponentially, illustrating the effects of structure loss (A) and retrieval slowing (B). τ represents mean latency; AD = Alzheimer's disease; NC = normal control.

Experiment 1

As just described, if AD results in the loss of interitem associations between the retrieval category cue and its corresponding exemplars, then AD participants should retrieve the available items more quickly, on average, than normal controls. By this same reasoning, young adult participants should retrieve category exemplars with a faster mean latency when the category is smaller. We tested this hypothesis in Experiment 1. To ensure that the smaller categories are indeed smaller, we made the small categories subsets of the larger categories (e.g., *farm animals* and *animals*).

A similar experiment was reported by Herrmann and Murray (1979), though there exist several differences between that experiment and this experiment. For example, participants in the present experiment responded aloud, whereas the participants in the Herrmann and Murray experiment wrote their responses on paper and drew a line underneath their most recent response every 15 s. Thus, our procedures provided a finer grained analysis of the time course of retrieval.

Method

Participants. Twelve undergraduates of the University of California, San Diego, participated for course credit.

Materials. The large categories and their associated smaller categories are listed in Appendix A. The category pairs were selected so that the small category contained less than a dozen exemplars. The same participant never received both a large category and its associated smaller category.

Design. Each participant completed 2 practice trials and 10 scored trials, 5 in each condition. Both the selection of these categories and their order of presentation were randomized for each participant.

Procedure. Participants were tested by computer in the presence of an experimenter. Each trial began with a brief warning tone and the screen-displayed prompt "Get ready to say examples of," which appeared for 4 s and was read aloud by the experimenter. The category name next appeared on the screen and remained for 60 s. Participants responded aloud, and a voice key recorded each response latency to an accuracy of 1 ms. Each response latency appeared on a computer screen (not visible to the participant) so that the experimenter could identify any response latency that was due to an extraneous noise such as a cough. Repetitions or incorrect responses were marked as such. A 20-s rest followed each trial.

Measure of latency for Experiments 1, 2, and 3. Two measures of response latency are presented in each of the three experiments in this article: *first-response latency* and *subsequent-response latency.* First-response latency measures the time until the first response, which includes the processing of the retrieval cue and the initiation of the search process. Subsequent-response latency represents the duration between the first response and each subsequent response. For example, if four responses in a particular trial occurred at 10, 11, 14, and 17 s, the first-response latency equals 10 s, and the mean subsequent-response latency equals the average of 1, 4, and 7 s or 4 s. In essence, subsequent-response latency equals mean latency once the retrieval process has begun. Subsequent-response latency is represented by τ, the estimate of mean latency discussed in the introduction, and serves as the measure of primary interest.

By excluding the initiation processes that precede the first response, subsequent-response latency provides a more accurate depiction of the retrieval process itself. In fact, the experimental manipulations presented herein affected differentially first- and subsequent-response latency, suggesting that the durations of the initiation stage and the search stage are indeed independent. However, our decision to focus primarily on estimates of subsequent-response latency rather than observed total latency (as measured since the onset of the recall period) did not affect the conclusions in this article. For each of the four comparisons of mean subsequent-response latency (τ) in this article, the corresponding pair of observed total latencies differed in the same direction. Moreover, three of the four comparisons of observed total latency produced a significant difference, with the fourth falling just short of significance ($p = .06$).

The best fitting exponentials were determined by least squares minimization. In each of the eight conditions, the two-parameter exponential accounted for a greater proportion of the variance than did a two-parameter line (averaged across conditions, .95 v .73). The asymptotic standard errors for each parameter were derived from the Hessian matrix of second partial derivatives (Maindonald, 1984). Pairwise comparisons of parameter values were performed by a t test. For these t tests, the asymptotic standard error of each parameter value provided a measure of the variability of each parameter, and the degrees of freedom for each of the two curve fits, summed together, provided the number of degrees of freedom (Ratkowski, 1983).

Results and Discussion

Not surprisingly, participants produced significantly more exemplars of large categories than small categories, $F(1, 11) = 95.91, p < .001$, as reported in Table 1. Errors (repetitions and false alarms) occurred less than once per hundred correct responses in both conditions and were excluded from the analyses.

Mean first-response latency in the two conditions, shown in Table 1, did not significantly differ, $F(1, 11) < 1$. Subsequent-response latencies were grouped into 5-s bins and plotted as a function of time (Figure 5). Specifically, each data point represents the average number of responses per trial that were produced in that 5-s bin. The best fitting two-parameter exponential accounted for a large portion of the variance in each condition (Table 1). The parameter estimate of asymptotic total (N) in the small-category condition (6.23) was close to the observed total less one (6.62), as would be expected given the removal of the first response. The estimate of N in the large-category condition (18.52) was greater than the observed response total less one (15.05), as retrieval had not yet reached the floor by the end of the recall period, as shown in Figure 5.

Of primary interest, the estimate of subsequent-response latency (τ) for the small-category condition was considerably less than that for the large-category condition, as shown in

Table 1
Response Totals and Response Latencies (in Seconds) for Experiment 1

Category size	Response total		First-response latency		Subsequent-response latency (τ)		
	M	*SE*	*M*	*SE*	*M*	*SE*	VAF
Small	7.62	0.46	3.62	0.56	11.68	1.27	.95
Large	16.05	1.03	4.30	0.97	32.93	1.16	.99

Note. Subsequent response latency excludes time until first response and is estimated by an exponential fit (see text). VAF denotes the variance accounted for by the exponential.

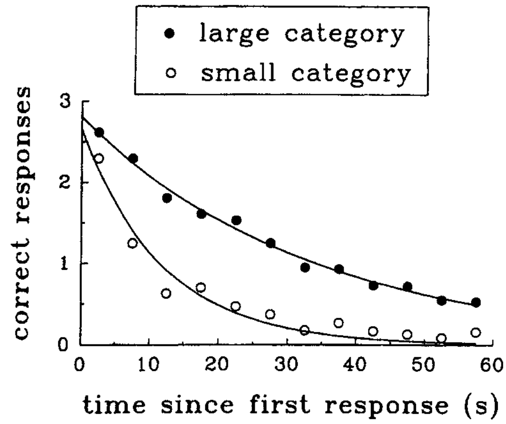

Figure 5. Mean recalled items for each 5-s bin and the best fitting exponentials for small and large category fluency in Experiment 1. Latency is measured from the first response.

Table 2, $t(10 + 10) = t(20) = 12.35, p < .001$. Consequently, the curves in Figure 5 reveal that the large-category half-life is roughly 20 s, whereas the small-category half-life is less than 10 s. Thus, retrieval in the small-category condition "approached the floor" more rapidly.

In summary, the retrieval of exemplars from small categories resulted in a faster mean latency. If the retrieval of exemplars from smaller categories is assumed to invoke mental search sets with fewer items, then mean latency is seen to vary directly with the number of these items. That mean latency depends on the size of the search set is also predicted by the random-search model, a scheme that is consistent with the observed exponential decline in retrieval seen in Experiment 1. Finally, if the structure-loss hypothesis is correct and AD participants are generating exemplars from smaller search sets, then the presence of Alzheimer's disease should produce faster mean latencies in the same manner as the use of smaller categories.

Experiment 2

As described in the introduction, the retrieval-slowing hypothesis suggests that the progression of Alzheimer's disease is accompanied by slower cognitive processing, which should increase the mean response latency in a category fluency task. As an illustration of this effect, in the second experiment we examined the category fluency of college-age participants while they manually responded to a continual and rapid repetition of a choice-response time task. Given that this secondary task is appropriately interfering, participants should produce category exemplars with a slower mean latency than when generating exemplars without the secondary task.

Such a finding would be consistent with the results of two experiments reported by Baddeley et al. (1984). In those experiments, participants maintained information within short-term store while either judging whether an item was an exemplar of a given category or judging whether a statement was true or false. Both judgments required information from *semantic memory,* and both were significantly slowed by the presence of the interfering task.

Method

Participants. Twelve undergraduates of the University of California, San Diego, participated for course credit.

Materials. The categories governing participants' responses are listed in Appendix B. To ensure that participants' recall totals would approach asymptote within 60 s even when performing the secondary task, we used relatively small categories in Experiment 2.

Design. Each participant completed 6 practice trials, 2 in the single-task condition and 4 in the dual-task condition, and 10 scored trials, 5 in each condition. Both the selection of these categories and their order of presentation were randomized for each participant.

Procedure. Except for the presence of the dual task in half of the trials, the procedure in Experiment 2 was identical to that of Experiment 1. In the dual-task trials, each trial began with a warning tone and the screen-displayed prompt "Get ready to TAP KEYS and SAY EXAMPLES of," which remained on the screen for 4 s. During that time participants placed the first three fingers of their left hand upon the 1, 2, and 3 keys of the keyboard numeric pad. As the category name appeared on the screen (and was then read aloud by the experimenter), three symbols simultaneously appeared on the computer screen, subtending approximately 10° of visual arc. One, two, or three of these symbols were bright red asterisks, and the remaining ones were dim blue dashes. Participants had to push the 1, 2, or 3 key, depending on whether the number of asterisks equaled one, two, or three, respectively. Once a key was pressed, the symbols disappeared from the screen. If an incorrect key was pressed, the word *ERROR* appeared in red and remained until the next presentation. The interstimulus interval was 800 ms. At the end of each trial, participants' response accuracy for that trial was displayed during the rest period. Participants were told that they should strive for "at least 90% accuracy" in the manual response task.

Results and Discussion

Not surprisingly, participants produced significantly fewer exemplars when performing the secondary task, $F(1, 11) = 6.39, p < .03$, as reported in Table 2. Errors (repetitions and false alarms) occurred less than once per hundred correct responses in both conditions and were excluded from the analyses. For the manual response task, mean accuracy equaled .85, and the mean response time of those correct responses equaled 512 ms.

Mean first-response latency in the single-task condition was significantly faster than that of the dual-task condition, $F(1, 11) = 27.0, p < .001$, as reported in Table 2. Presumably, the initiation of the search process was slowed by the secondary task.

Table 2
Response Totals and Response Latencies (in Seconds) for Experiment 2

Condition	Response total *M*	Response total *SE*	First-response latency *M*	First-response latency *SE*	Subsequent-response latency (τ) *M*	Subsequent-response latency (τ) *SE*	VAF
Single task	4.83	0.21	3.45	0.28	4.40	0.18	.99
Dual task	3.88	0.33	6.27	2.65	8.07	0.61	.97

Note. Subsequent response latency excludes time until first response and is estimated by an exponential fit (see text). VAF denotes the variance accounted for by the exponential.

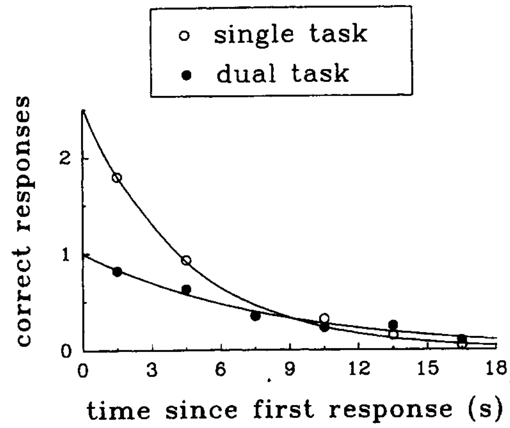

Figure 6. Mean recalled items for each 5-s bin and the best fitting exponentials for the single and dual task conditions in Experiment 2. Latency is measured from the first response.

As before, subsequent-response latencies were grouped into bins and plotted as a function of time, as shown in Figure 6. Because participants produced more than half of their responses within 5 s of the first response, 3-s bins were used instead of 5-s, but bin size has very little effect on parameter estimates. Though Figure 6 includes only the first portion of the recall period (because almost all of the responses occurred early on), the exponential was fit to the data for the entire 60 s. The exponential again provided a smooth visual fit (Figure 6) and accounted for a large portion of the variance in each condition (Table 2). The parameter estimates of asymptotic recall total (N) in both the single-task (3.71) and the dual-task condition (2.69) were very close to the corresponding observed totals less one (3.83 and 2.88, respectively) because retrieval in both conditions had reached the floor well before 60 s, as shown in Figure 6.

As reported in Table 2, mean subsequent-response latency (τ) in the dual-task condition was significantly slower than that in the single-task condition, $t(18 + 18) = t(36) = 5.77, p < .001$. This effect is also seen in Figure 6, as the response latencies in the dual-task condition are more evenly distributed across the recall period. Thus, the dual-task function declines by one half in about 6 s, whereas the single-task function declines by one half in about 3 s.

In summary, the retrieval of exemplars while performing a secondary task resulted in a slower mean response latency. If a secondary task is assumed to slow the retrieval from the mental search set, then mean latency is seen to depend directly on processing speed. That mean latency depends on processing speed is also predicted by the random-search model, a scheme that is consistent with the exponential decline in retrieval. Finally, if the retrieval-slowing hypothesis of Alzheimer's disease is correct and AD participants are simply slowed down, then the presence of Alzheimer's disease should produce slower mean latencies in the same manner as a secondary task.

Notably, the secondary task reduced the number of responses but also increased the mean latency of those responses (thereby producing the intersecting exponentials seen in Figure 6). No previous semantic memory experiment with either patients or healthy participants has produced this inverse variation between N and τ. Instead, all such studies have yielded a direct variation, as was observed in Experiment 1. Taken together, Experiments 1 and 2 illustrate the independence of N and τ, as described in more detail in the General Discussion section.

Though the secondary task reduced response totals, the effect was relatively small. Likewise, Craik, Govoni, and Naveh-Benjamin (1993) reported a finding in which the presence of a secondary choice-response time task during retrieval of study list items produced only a slight reduction in the response total. Given the large effect of the secondary task on response latency in Experiment 2, it may be the case that secondary demands have little effect on the number of available items within the search set, yet they greatly reduce the speed at which these items can be retrieved.

Experiment 3

In this experiment we examined the response latencies of both AD and NC participants during category and letter fluency tasks. Mean latency, once retrieval is underway, is again the measure of primary interest. As delineated in the introduction, the structure-loss hypothesis predicts that AD participants will produce category exemplars with a faster mean latency than controls, as illustrated by the small-category condition in Experiment 1. In contrast, the retrieval-slowing hypothesis predicts that AD participants will produce category exemplars with a slower mean latency than controls, as illustrated by the dual-task condition in Experiment 2.

Method

Participants. Participants included 12 outpatients with a diagnosis of probable Alzheimer's disease and 13 normal controls, all of whom were tested and diagnosed at the Alzheimer's Disease Research Center of the University of California, San Diego. The testing of 1 additional AD participant was discontinued after he became very distraught. All participants were native English speakers. Written informed consent was obtained from each participant or his or her caregiver.

The diagnosis of Alzheimer's disease was given by two senior staff neurologists according to both the criteria for primary degenerative dementia put forth in the *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., rev.) by the American Psychiatric Association (1980) and the criteria for probable Alzheimer's disease as established by the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association task force (McKhann et al., 1984). In addition, other possible causes of dementia were ruled out by medical, laboratory, and neuropsychological testing. Both neurologists were unaware of the participants' performance in this experiment. NC participants were either volunteers recruited through newspaper advertisements or spouses of patients. NC participants with a learning disability, serious neurologic or psychiatric illness, or history of alcohol or drug abuse were excluded.

Table 3 includes the demographic and diagnostic characteristics of the AD and NC participants. The two groups did not differ with respect to gender composition ($\chi^2 < 1$), age ($t < 1$), or education ($t < 1$). The two groups did significantly differ with respect to Dementia Rating Scale (DRS) scores, $t(23) = 42.20, p < .001$, and Mini-Mental State Examination (MMSE) scores, $t(23) = 15.99, p <$

Table 3
Participants' Demographic and Diagnostic Characteristics for Experiment 3

Characteristic	Alzheimer's group	Control group
Sex		
Male	7	7
Female	5	6
Age (in years)		
M	75.1	74.6
SE	1.5	2.0
Range	67–84	62–87
Education (in years)		
M	14.3	14.5
SE	1.1	0.9
Range	8–20	8–19
DRS score		
M	99.6	139.4
SE	4.6	1.4
Range	68–122	130–144
MMSE		
M	18.3	29.2
SE	1.4	0.3
Range	12–25	26–30

Note. DRS = Dementia Rating Scale; MMSE = Mini-Mental State Examination.

.001. Three of the AD participants had DRS scores greater than 115 and MMSE scores greater than 20, thereby qualifying them as "mild."

Materials. Appendix C includes the category and letter cues governing the participants' responses. These categories were used in the large-category condition in Experiment 1, with three exceptions. *Fruits* and *vegetables* replaced *food* and *European countries* replaced *countries* so that the categories would be more similar in size. *Presidents* replaced *sports* because previous researchers have reported a gender difference in older participants when generating sports.

Design. Each participant underwent 2 practice trials (one-category fluency and one-letter fluency) and 16 scored trials with alternating category and letter trials. Each participant received the same eight categories and eight letters, but the ordering of these categories and letters was randomly chosen for each participant to eliminate order effects.

Procedure. Participants were tested by computer in the presence of an experimenter. The nature of the task was explained to participants, all of whom had previously participated in verbal fluency tasks while undergoing diagnostic testing. Participants were instructed to begin as soon as the retrieval cue appeared on the screen. Participants were instructed to avoid proper nouns (Tom) and variants (run, running, runner, etc.) during the letter fluency trials.

Each trial began with a brief warning tone and the screen-displayed prompt "Say examples of" or "Say words that begin with the letter." The prompt appeared for 4 s and was read aloud by the experimenter. The category or letter cue next appeared on the screen and was immediately read aloud by the experimenter. The category name or letter remained on the screen for 60 s. A 30-s rest period followed each trial.

The experimenter recorded each response with an immediate key tap. Though this technique does not provide the millisecond accuracy of a voice key, the lag between the participant's voice response and the experimenter's manual response is relatively consistent and lasts only a couple-hundred milliseconds. Moreover, because response latencies were grouped into 5-s bins, this loss of temporal resolution is effectively inconsequential. Each session was tape recorded.

Results and Discussion

After testing only a few participants, we decided that the category of *U.S. states* should be excluded from the analyses for two reasons. First, participants in these trials produced exemplars such that consecutively generated states were almost always contiguous states (e.g., *Nevada, Idaho, Wyoming, Colorado,* and so forth). Such a pattern suggests that these items were retrieved by the visualization of a map, which, as Indow and Togano (1970) found, produces qualitatively different results. Second, the number of exemplars in this condition given by these first few participants was two to three times that given in other categories, thereby greatly reducing the homogeneity of the data. More problematically, however, these initial participants were still producing exemplars of *states* at a considerable rate at the end of the recall period, suggesting that the recall period was much too short for a category of this size.

The responses from the other categories were pooled together, as were those given for each letter cue. It has previously been reported that the category fluency of AD participants is consistently impaired across a wide range of categories (Chertkow & Bub, 1990).

Errors. The number of extralist intrusions and repetitions given by AD and NC participants are reported in Table 4. AD participants produced more intrusions, though only the difference in the category task was significant, $F(1, 23) = 6.90, p < .02$. AD participants also produced more repetitions than controls in both fluency tasks, though both of these differences fell just short of significance. Two-way analyses of variance (ANOVAs) yielded nonsignificant interactions between participant group and fluency task for both intrusions and repetitions.

Correct responses. As reported in Table 5, AD participants produced significantly fewer correct responses than controls in both the category condition, $F(1, 23) = 84.09, p < .001$, and the letter condition, $F(1, 23) = 23.83, p < .001$. A two-factor ANOVA (Group × Condition) revealed not only the expected significant main effect of group, $F(1, 23) = 54.07, p < .001$, but also a Participant Group × Fluency Task interaction, $F(1, 23) = 6.15, p = .02$. Thus, in terms of the number of correct responses, AD participants exhibited greater relative impairment in the category condition. Given that category fluency relied more heavily on interitem associations than does letter fluency, this result is consistent with the hypothesis that the

Table 4
Errors in Experiment 3

Participants	Extralist intrusions *M*	Extralist intrusions *SE*	Response repetitions *M*	Response repetitions *SE*
	Category fluency			
Alzheimer's	0.37	0.09	0.46	0.12
Control	0.11	0.05	0.20	0.07
	Letter fluency			
Alzheimer's	0.19	0.06	0.59	0.14
Control	0.08	0.03	0.32	0.12

Table 5
Response Totals and Response Latencies (in Seconds) for Experiment 3

Participants	Correct response total *M*	*SE*	First-response latency *M*	*SE*	Subsequent-response latency (τ) *M*	*SE*	VAF
			Category fluency				
Alzheimer's	4.10	0.48	11.33	1.74	16.89	2.41	.89
Control	13.82	0.92	5.79	0.38	27.37	1.45	.98
			Letter fluency				
Alzheimer's	6.92	1.01	7.47	0.79	49.10	5.49	.89
Control	13.93	1.02	4.70	0.37	49.96	3.56	.96

Note. Subsequent response latency excludes time until first response and is estimated by an exponential fit (see text). VAF denotes the variance accounted for by the exponential.

semantic memory impairments in AD participants were due to the loss of interitem associations.

As in Experiments 1–2, we separately analyzed first-response latency and subsequent-response latency to obtain a much truer picture of the retrieval dynamics. The mean first-response latency of AD participants was significantly slower than that of controls in both the category condition, $F(1, 23) = 10.46, p < .01$, and the letter condition, $F(1, 23) = 10.70, p < .01$, as reported in Table 5.

Subsequent-response latencies were again grouped into 5-s bins and plotted as a function of time since the first response, as shown in Figure 7. The best fitting exponentials accounted for a large portion of the variance in these data in each condition (Table 5). In the category condition, estimates of asymptotic recall (*N*; 3.61 for AD participants and 14.75 for NC participants) were greater than the corresponding observed totals less one (3.10 for AD participants and 12.82 for NC participants). This difference was relatively small for the AD participants because their retrieval was close to the floor after 60 s, as shown in Figure 7. In the letter fluency condition, estimates of *N* (9.71 for AD participants and 19.19 for NC participants) were greater than the corresponding observed totals less one (5.92 for AD participants and 12.92 for NC participants), as neither group of participants had finished responding before 60 s.

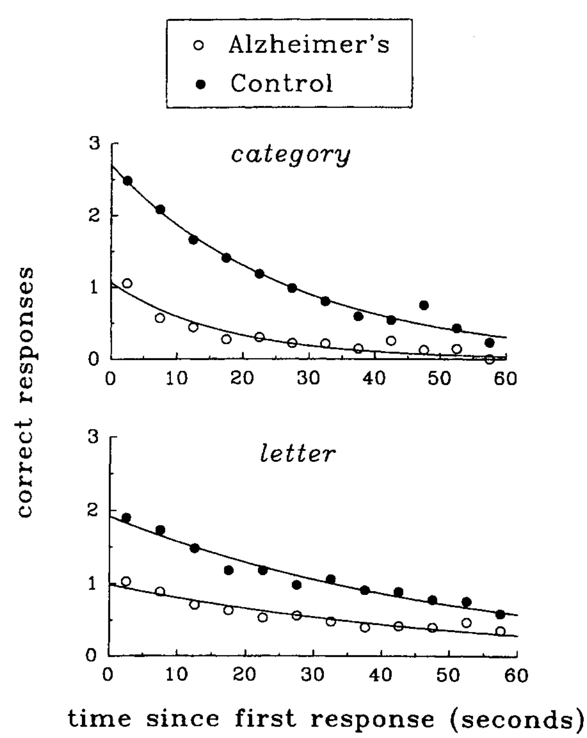

Figure 7. Mean recalled items for each 5-s bin in Experiment 3. Latency is measured from the first response.

With regard to subsequent-response latency, AD participants differed from controls in the category condition only (Table 5). For the category condition, the estimate of subsequent-response latency (τ) for AD participants was significantly less than that of controls, $t(10 + 10) = t(20) = 3.73, p < .001$. In the letter condition, however, values of τ were almost identical ($t < 1$). Again, these differences are apparent in the plots of Figure 7. In the category condition, retrieval by AD participants declined by one half in about 10 s, whereas that by NC participants declined by one half in about 18 s. In the letter condition, retrieval by both groups of participants declined by one half in about 35 s. Finally, it is worth noting that, in the category condition, AD participants produced not only a faster subsequent-response latency (τ) but a faster total mean latency as well. Thus, even though AD participants were much slower to begin, their observed total latency, measured from the onset of the recall period, was still faster than that of NC participants. In summary, the reduced mean latency for AD participants in the category fluency task is consistent with the view that AD causes the loss of interitem associations within semantic memory.

General Discussion

In three experiments, response latency was measured while participants generated exemplars of presented categories. In Experiment 1, the use of smaller categories decreased mean response latency (as measured from the onset of responding). In Experiment 2, a concurrent secondary task increased mean response latency. In Experiment 3, the mean response latency of AD participants was faster than that of elderly controls in the category fluency task but not in the letter fluency task. However, AD participants were slower than controls in getting started in both tasks, as measured by first-response latency.

That AD participants produced category exemplars with a faster mean latency than controls is theoretically consistent with the view that Alzheimer's disease is characterized by the loss of interitem associations within semantic memory. Specifically, if the hierarchical associations between a category name and its exemplars (or the associations between exemplars) are lost, the activation of that category name as a retrieval cue will result in the activation of fewer category exemplars. With fewer available category exemplars within the search set, less time is needed to find the items, and, consequently, mean latency quickens. Indeed, this is exactly what happened in Experiment 1—retrieval cues that subsumed a small number of items (i.e., the smaller categories) produced significantly faster mean latencies.

Also consistent with the structure-loss hypothesis is the observed interaction between task (category or letter) and participant group (AD or NC) for both mean latency and response total in Experiment 3. That is, if it is assumed that category fluency depends on interitem associations to a greater extent than does letter fluency, the loss of interitem associations should impair performance in the category fluency task to a greater extent than in the letter fluency task. Such an assumption seems particularly valid in light of the necessity of semantic interitem associations. Whereas semantic similarity is the cornerstone of abstract thought, phonemic similarity seems to serve no adaptive function.

Rival Hypotheses

The primary finding of the present article, the reduced mean latency of category exemplars produced by AD participants, can be explained a posteriori by processes other than structure loss. For instance, AD participants may have terminated their search prematurely, thereby producing artifactually few items in the later portion of the recall period and a consequently decreased mean response latency. For example, AD participants may have simply been unable or unwilling to stay focused on the task at hand. Similarly, it could be argued that AD participants used a much stricter stopping rule than did the NC participants. That is, after relatively few retrieval failures by the search mechanism, the AD participants ceased retrieval.

Though these two accounts cannot be ruled out definitively, neither is consistent with a secondary analysis of the present data. Additional estimates of τ_{AD} and τ_{NC} for the category fluency condition were obtained for each of three truncated recall periods: 45 s, 30 s, and 15 s. These estimates, along with the original estimates of τ_{AD} and τ_{NC} for the entire 60 s, are shown in Table 6. If AD participants terminated their search early, for any reason, the difference between τ_{AD} and τ_{NC} should decrease as the effective recall period is shortened. Instead, this difference actually increased with each successive shortening of the recall period, as τ_{NC} remained roughly constant, whereas τ_{AD} decreased.

Alternatively, the reduced mean latency of AD participants may have resulted from greater-than-normal output interference. Findings reported by Blaxton and Neely (1983), for example, illustrate that the generation of several category exemplars can inhibit the subsequent generation of exemplars from the same category. Thus, Alzheimer's disease may be characterized by greater-than-normal category-specific inhibition. However, not every instantiation of output interference can account for the findings reported in this article. Rundus (1973) and Shiffrin (1970), for example, proposed that each sampling of an item increases the strength of that item and, therefore, decreases the probability that other not-yet-sampled items will be retrieved. According to this interpretation, increased output interference delays the retrieval of some items so that mean latency is increased, not decreased. However, if a stopping rule is assumed to exist, then this characterization of output interference can, under certain conditions, account for the present findings. Indeed, simulations performed in our laboratory reveal that a combination of a faster-than-normal stopping rule and greater-than-normal strengthening of previously retrieved items can produce the findings discussed earlier, including the positive correlation between τ_{AD} and recall period duration shown in Table 6.

Nevertheless, if increased category-specific inhibition in AD participants is to account for the sizable difference between τ_{AD} and τ_{NC} reported in this article, it would need to be extreme. In fact, it might be better depicted as an underlying perseveration in the retrieval process. Of course, such an explanation of the present data is post hoc and should be tested in future research. For example, the presentation of a few category exemplars as cues for generating the remaining category exemplars might produce a greater reduction in response totals for AD participants than for NC participants. This finding of a greater-than-normal part-list cuing effect for AD participants would be consistent with a retrieval-perseveration model.

Accounting for Nebes's Data

That AD participants generated category exemplars with a faster mean latency than controls is, of course, not consistent with the retrieval-slowing hypothesis. If the impaired semantic memory of AD patients resulted from slowed processing, then the average time to find an item should increase. As was illustrated in Experiment 2, slower processing (as was induced by a concurrent secondary task) increases mean latency.

Yet the results reported in this article should not be construed as evidence against slowed processing in AD patients; but they should instead be construed as evidence against the hypothesis that this slowed processing is the primary cause of the semantic memory impairments seen in AD patients. The work by Nebes and his colleagues leaves little doubt that AD patients are slower in a wide range of cognitive tasks. Indeed, AD participants in our study were slower in initiating the search process, as measured by first-response latencies. It is, therefore, reasonable to believe that AD patients are also slower during the search process. However, the evidence presented herein strongly suggests that the impaired word-finding ability of AD patients stems primarily from the loss of semantic structure. That is, it may be the case that slower processing contributes, in part, to the word-finding difficulties of AD patients, but the loss of interitem associations within semantic memory is the predominant

Table 6
Subsequent-Response Latencies (τ; in Seconds) for Category Fluency as a Function of Recall Period Duration for Experiment 3

Time since first response	τ_{NC} *M*	τ_{NC} *SE*	τ_{AD} *M*	τ_{AD} *SE*	$\tau_{NC} - \tau_{AD}$
60 s	27.37	1.45	16.89	2.41	10.48
45 s	26.03	0.50	15.98	2.63	10.05
30 s	26.54	0.59	13.08	1.96	13.44
15 s	27.61	3.53	10.26	2.10	17.35

Note. Standard errors are asymptotic. NC = normal control; AD = Alzheimer's disease.

cause. However, if slowed processing did characterize the AD participants in the present study, then it worked against the reported finding of their faster mean latency.

Recall, however, that the data put forth by Nebes and his associates were consistent with both retrieval slowing and intact semantic structure. As described in the introduction, Nebes measured response times across conditions of varying semantic association and found that the response times of AD and NC participants were affected in a similar manner. Consider, for example, the category decision task in Nebes et al. (1986), as it is most similar to the category fluency task. Participants were audibly presented with a semantic category (e.g., *bird*) and then visually presented with an item that was or was not an exemplar of the category. For the yes trials, category exemplars were either strong associates (e.g., *robin*) or weak associates (e.g., *penguin*). Though NC participants were faster than AD participants in both conditions, the response times of both AD and NC participants were about 105 ms faster for strong associates than for weak associates. That is, there was a main effect of both participant groups and associate strength, but no interaction. As interpreted by Nebes et al. (1986), the "category name primed recognition . . . to the same degree in normal and demented subjects" (p. 268). However, even though both AD and NC participants exhibited the same absolute decrease in response time, the NC participants exhibited a greater relative decrease. Moreover, the difference in the relative decreases grows more salient in view of the fact that response times can only decrease so much before encountering a floor. That is, the 105-ms priming effect for NC participants may have reduced their response time to a theoretical minimum equal to, say, the time needed to just perceive the stimulus and execute the response. (Granted, it could be argued that Alzheimer's disease slows stimulus perception and response execution while leaving retrieval speed intact, but that argument seems untenable.) Finally, it is worthy of mention that AD participants did make more errors than NC participants in these experiments. That in itself would seem to argue against an intact semantic structure.

Implications for Memory-Intact Individuals

Though the results of Experiments 1 and 2 provide an illustration of the structure-loss and retrieval-slowing hypotheses, these findings are important in their own right. First, these findings provide direct evidence for the dependence of mean latency on both search-set size and processing speed—the central tenet of the random-search model. Likewise, the decline in rate of retrieval in all three experiments was well described by the exponential, which is also consistent with the random-search model.

Second, the results of Experiments 1 and 2 provide an important dissociation of response total and response latency. Whereas the manipulation in Experiment 1 and all previous category fluency experiments produced a direct variation between the number of items recalled (N) and their mean latency (τ), the presence or absence of dual-task interference in Experiment 2 resulted in an inverse variation in these two parameters. Specifically, the interference decreased response total and increased mean latency. The finding serves to dissociate these two dependent measures, thereby ruling out any inherent relationship between them. For example, suppose that longer mean latencies simply derived from the greater amount of time needed to produce more responses aloud. If so, then any manipulation that increased recall total would also increase recall latency. However, the dissociation of recall total and recall latency provided by Experiments 1 and 2 rules out any such a priori relationship. (Incidentally, the time between responses is generally much greater than the vocal duration of each response, further arguing against the notion that responses are queueing up and awaiting motor execution.)

This dissociation of recall total and recall latency has also been reported in the free recall of items from episodic memory. For example, longer study lists result in both longer mean latencies and greater recall totals (Rohrer & Wixted, 1994), whereas the build-up of proactive interference results in longer mean latencies but smaller recall totals (Wixted & Rohrer, 1993). Notably, both results are consistent with the conceptualization of mean latency as a measure of the size of the search set, as it is commonly believed that larger mental search sets result after studying longer lists or studying several lists of words belonging to the same category. Thus, as the retrieval cue subsumes a greater number of items, more time is required, on average, to retrieve those items.

Future Directions

Though the results presented in this article suggest that retrieval slowing cannot account for the word-finding impairments in AD patients, evidence suggests that retrieval-slowing may be responsible for the retrieval deficits exhibited by HD patients. For instance, HD patients produce significantly fewer items than controls in both category and letter fluency tasks (Butters et al., 1987; Monsch et al., 1994). If Huntington's disease does result in retrieval slowing, then these patients should retrieve category exemplars with a slower mean latency than controls. If this result were found, then mean latency would enjoy a rare distinction as a dependent variable. Namely, mean latency would be decreased in one patient group (AD) and increased in another (HD). Furthermore, because both patient groups produce fewer responses than controls, such a finding would further illustrate the importance of response latency analyses in studies of free recall from either episodic or semantic memory. Such analyses divulge the most salient feature of free recall—its time course—and are perhaps best suited for revealing the underlying processes of retrieval.

References

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.

Baddeley, A., Lewis, V., Eldridge, M., & Thomson, N. (1984). Attention and retrieval from long-term memory. *Journal of Experimental Psychology: General, 113,* 518–540.

Blaxton, T. A., & Neely, J. H. (1983). Inhibition from semantically related primes: Evidence of a category-specific inhibition. *Memory & Cognition, 11,* 500–510.

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.

Butters, N., Granholm, E., Salmon, D. P., Grant, I., & Wolfe, J. (1987). Episodic and semantic memory: A comparison of amnesic and demented patients. *Journal of Clinical and Experimental Neuropsychology, 9,* 479–497.

Chertkow, H., & Bub, D. (1990). Semantic memory loss in dementia of the Alzheimer's type. *Brain, 113,* 397–417.

Craik, F. I. M., Govoni, R., & Naveh-Benjamin, M. (1993, November). *The effects of divided attention at encoding and retrieval.* Paper presented at the 34th annual meeting of the Psychonomic Society, Washington, DC.

Graesser, A., & Mandler, G. (1978). Limited processing capacity constrains the storage of unrelated sets of words and retrieval from natural categories. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 86–100.

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

Gruenewald, P. J., & Lockhead, G. R. (1980). The free recall of category examples. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 225–240.

Herrmann, D. J., & Chaffin, R. J. S. (1976). Number of available associations and rate of association for categories in semantic memory. *Journal of General Psychology, 95,* 227–231.

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.

Herrmann, D. J., & Pearle, P. M. (1981). The proper role of clusters in mathematical models of continuous recall. *Journal of Mathematical Psychology, 24,* 139–162.

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

Johnson, D. M., Johnson, R. C., & Mark, A. L. (1951). A mathematical analysis of verbal fluency. *Journal of General Psychology, 44,* 121–128.

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

Martin, A., & Fedio, P. (1983). Word production and comprehension in Alzheimer's disease: The breakdown of semantic knowledge. *Brain and Language, 19,* 124–141.

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.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease. *Neurology, 34,* 939–944.

Metlay, W., Handley, A., & Kaplan, I. T. (1971). Memory search through categories of varying size. *Journal of Experimental Psychology, 91,* 215–219.

Mickanin, J., Grossman, M., Onishi, K., Auriacombe, S., & Clark, C. (1994). Verbal and nonverbal fluency in patients with probable Alzheimer's disease. *Neuropsychology, 8,* 385–394.

Monsch, A. U., Bondi, M. W., Butters, N., Paulsen, J. S., Salmon, D. P., Brugger, P., & Swenson, M. R. (1994). A comparison of category and letter fluency in Alzheimer's and Huntington's disease. *Neuropsychology, 8,* 25–30.

Morrison, D. G. (1979). An individual differences pure extinction process. *Journal of Mathematical Psychology, 19,* 307–315.

Nebes, R. D., Boller, F., & Holland, A. (1986). Use of semantic context by patients with Alzheimer's disease. *Psychology and Aging, 1,* 261–269.

Nebes, R. D., & Brady, C. B. (1990). Preserved organization of semantic attributes in Alzheimer's disease. *Psychology and Aging, 5,* 574–579.

Nebes, R. D., & Brady, C. B. (1991). The effect of contextual constraint on semantic judgments by Alzheimer patients. *Cortex, 27,* 237–246.

Nebes, R. D., & Brady, C. B. (1992). Generalized cognitive slowing and severity of dementia in Alzheimer's disease: Implications for the interpretation of response-time data. *Journal of Clinical and Experimental Psychology, 14,* 317–326.

Nebes, R. D., Brady, C. B., & Huff, F. J. (1989). Automatic and attentional mechanisms of semantic priming in Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology, 11,* 219–230.

Nebes, R. D., Martin, D. C., & Horn, L. C. (1984). Sparing of semantic memory in Alzheimer's disease. *Journal of Abnormal Psychology, 93,* 321–330.

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. (1964). Composition of associative clusters. *Journal of Experimental Psychology, 67,* 199–208.

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., & Payne, D. G. (1985). Response criteria do not affect recall level or hypermnesia: A puzzle for generate-recognize theories. *Memory & Cognition, 13,* 1–7.

Roediger, H. L., III, Payne, D. G., Gillespie, G. L., & Lean, D. (1982). Hypermnesia as determined by level of recall. *Journal of Verbal Learning and Verbal Behavior, 21,* 635–655.

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

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

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

Rundus, D. (1973). Negative effects of using list items as recall cues. *Journal of Verbal Learning and Verbal Behavior, 12,* 43–50.

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

Tröster, A. I., Salmon, D. P., McCullough, D., & Butters, N. (1989). A comparison of the category fluency deficits associated with Alzheimer's and Huntington's disease. *Brain and Language, 37,* 500–513.

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., & Rohrer, D. (1993). Proactive interference and the dynamics of retrieval. *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 literature. *Psychonomic Bulletin and Review, 1,* 89–106.

Appendix A
Categories Used in Experiment 1

Large category	Small category
Musical instruments	Stringed instruments
Domesticated animals	Farm animals
Wild animals	Primates
Bodies of salt water	Oceans
Bodies of fresh water	Lakes
Sports	Team sports
Countries	South American countries
Occupations	Occupations requiring advanced degrees
Foods	Breakfast foods
States of the United States	Coastal states of the United States

Appendix B
Categories Used in Experiment 2

Category
Planets
Branches of the military
Colleges of the University of California, San Diego
Stringed instruments in orchestra
University of California campuses
Brass horns
The senses
Continents
Colors of the rainbow
Oceans

Appendix C
Categories and Letters Used in Experiment 3

Category	Letter
Fruits	A
Vegetables	C
Pets and farm animals	D
U.S. presidents	F
Countries in Europe	L
Musical instruments	M
Wild animals	P
U.S. States	S

Received April 21, 1994
Revision received August 29, 1994
Accepted September 6, 1994 ■