Does Expanded Retrieval Produce Benefits Over Equal-Interval Spacing? Explorations of Spacing Effects in Healthy Aging and Early Stage Alzheimer's Disease

David A. Balota, Janet M. Duchek, Susan D. Sergent-Marshall, and Henry L. Roediger III Washington University

Three experiments explored different schedules of retrieval practice in young adults, older adults, and individuals with dementia of the Alzheimer type. In each experiment, an initial acquisition phase was presented in which participants studied or attempted to retrieve response words to cues, followed by a later cued-recall test. Experiment 1 produced a benefit of expanded retrieval over equal-interval retrieval during acquisition, but this benefit was lost in final cued recall. In Experiments 2 and 3, participants received corrective feedback during acquisition and modified spacing schedules. There was again no evidence of a difference between expanded and equal-interval conditions in final cued recall. Discussion focuses on the potential benefits and costs of expanded retrieval on a theoretical and applied level.

Keywords: aging, Alzheimer's disease, memory, retrieval

The advantage of spaced learning episodes over massed learning episodes is one of the most fundamental characteristics in learning (see, e.g., Crowder, 1976, Chapter 9). In the present study, we examine an extension of the spacing effect that has considerable potential for maximizing the benefits of spacing. Specifically, a particular type of spacing that expands the interval between learning episodes appears to afford an additional advantage over a spacing schedule that equally spaces the study episodes. For example, in learning the name of a new person just introduced, one might use an expanded practice schedule by retrieving the name at an interval of 15 s, 45 s, and 60 s, compared with an equally spaced practice of retrieving the name pair every 40 s. This particular type of spacing, referred to as expanded retrieval practice, has been extended to applied situations, including educational settings (e.g., Rea & Modigliani, 1985), and to individuals with memory impairments, such as individuals with early stage dementia of the Alzheimer's type (DAT; e.g., Camp, Foss, Stevens, & O'Hanlon,

The goal of the present series of experiments is to explore the benefits of expanded retrieval in young adults, healthy older adults, and individuals with early stage DAT. First, we briefly

David A. Balota, Janet M. Duchek, Susan D. Sergent-Marshall, and Henry L. Roediger III, Department of Psychology, Washington University. This work was supported by National Institute on Aging Grants PO1 AGO3991 and P50AGO5681. We thank Martha Storandt for providing the psychometric test support, as well as John Morris and the clinicians at the Washington University Alzheimer's Disease Research Center for their careful recruitment and description of the healthy older adult and dementia of the Alzheimer's type participant groups. We also thank Dori Sabata for her help with study list construction and pilot testing, Priyank Shah for his help with Experiment 3, and Jes Logan for her help at various stages of this project.

Correspondence concerning this article should be addressed to David A. Balota, Department of Psychology, Washington University, St. Louis, MO 63130. E-mail: dbalota@artsci.wustl.edu

review the empirical evidence for this phenomenon, the potential theoretical accounts, and why one might be especially interested in this procedure in these particular groups of individuals.

Evidence for Benefits of Expanded Over Equal-Interval Spaced Retrieval

Landauer and Bjork (1978) were the first to fully explore the benefits of expanded retrieval over equal-interval spacing in two controlled experiments. In their first experiment, participants were initially presented a deck of cards that included fictitious first and second names of individuals. We refer to this as the *acquisition* phase. For present purposes, we focus on the trials in which participants were first presented both the first and second names intact, followed at varying schedules of receiving cards with only the first name of the study pair as a retrieval cue for the second name. For example, in the expanded retrieval schedule participants may receive the first name as a retrieval cue at 1, 4, and 10 intervening cards. This is in contrast to an equal-interval condition, in which the first name was presented as a retrieval cue after 5, 5, and 5 intervening cards. Hence, the average number of intervening items is equated across these two conditions. They used a massed condition (study of the name followed by three immediate tests) to obtain a baseline estimate of the influence of simple repetition without spacing. The results of a later cued-recall test yielded a large benefit for both the expanded and equal-interval condition over the massed condition. More important, there was a small but significant effect (approximately a 7% benefit, on the basis of interpolation from their Figure 2) of the expanded condition over the equal-interval condition in the final cued-recall test.

Cull, Shaughnessy, and Zechmeister (1996) noted that there has been surprisingly little published work on expanded retrieval practice, even though it would appear to have considerable potential in applied situations. They speculated that this possibly was due to the fact that the original observation appeared in a chapter. Cull et al. further explored the benefits of expanded retrieval over equal

spacing in a series of five experiments. They compared the expanded schedules of 1, 5, and 9 with the equal-interval schedule of 5, 5, and 5. In addition, like Landauer and Bjork (1978), they included a 0-0-0, massed presentation condition, as a baseline condition to measure the mere effect of repetition. In general, the results from their experiments supported the utility of expanded retrieval. We present their data in the top portion of Table 1. Experiments 1 and 4 produced reliable expanded retrieval benefits compared with an equal-interval condition. The results from Experiments 2 and 3 were in the same direction but did not reach significance, whereas Experiment 5 produced near-ceiling effects in both spaced conditions. In the latter experiment, participants were allowed to restudy the items during initial acquisition phase if they were unable to recall the target; hence, this condition is called test-study. Somewhat surprising, as shown in Table 1, the comparison of the equal-interval condition with the massed condition did not consistently produce spacing effects. A test for the spacing effect was not directly provided for Experiments 1 and 2, and in Experiments 3 and 4, these two conditions produced identical performances. Only their fifth experiment provided strong evidence for a spacing effect, but this experiment did not provide evidence for an expanded retrieval effect, because of ceiling problems. Hence, the lack of a spacing effect in two of the experiments, the lack of a reliable difference between the expanded and equalinterval condition in two of the experiments, and the potential ceiling effect in the fifth experiment restrict any strong inferences regarding the benefits of expansion from this study.

Cull (2000) provided a follow-up study to the initial Cull et al. (1996) study. In Experiment 1, which is the most relevant to the present discussion, Cull included the same three spacing schedules used in the earlier Cull et al. (1996) study, along with three different types of study. In the study-only condition, the cue and response item of a paired associate were both presented on each repetition for 8 s. In the test-only condition, the cue for retrieval was presented on each occasion for 8 s. In the test-study condition, the cue was presented for 6 s for retrieval, followed by the response item for 2 s. As shown in the bottom half of Table 1, Cull did not obtain an advantage for the expanded over the equal-interval condition in any of these conditions. In three follow-up

experiments, Cull also explored the influence of expanded versus uniform spacing under conditions that allow reviewing of the materials, and also with text materials, and again did not find evidence of better performance for an expanded spacing schedule compared with a uniform spacing condition. Hence, although spacing clearly affords a benefit, Cull questioned the utility of expanded retrieval over a matched equal-interval condition.

Most recently, Carpenter and DeLosh (2005, Experiment 2) have investigated face–name learning under massed (0–0–0), expanded (1–3–5), and matched equal-interval (3–3–3) conditions. Carpenter and DeLosh included a study procedure, as well as a study and test procedure, during the acquisition phase. As shown in Table 1, Carpenter and DeLosh found a large effect of spacing but no evidence of an additional benefit of expanded over equal-interval practice in either the study or the study and test procedures. In fact, there was a small reliable advantage for the equal-interval condition over the expanded condition in the final cued-recall test

Hence, we are aware of four studies that have been published with the necessary comparisons allowing one to explore the effects of expanded retrieval compared with equal-spaced retrieval and spaced retrieval compared with massed retrieval. Although the results from these studies initially yielded some evidence for a benefit of expanded over matched equal-interval conditions, the effect does not appear to be large and is clearly inconsistent across studies. Of course, in order for a strong case for the benefits of expanded spaced condition to be made, one would expect this condition to produce higher performance than the equally spaced condition, which in turn would produce the standard spacing effect.

Why Might Expanded Retrieval Practice Produce a Benefit Over an Equal-Interval Schedule?

Unfortunately, there has also not been much theoretical work addressing the mechanisms underlying the expanded retrieval effect. One account is based on the benefits of retrieval under easy and difficult retrieval situations by Bjork (1994). Bjork has noted that difficulties during study may force participants to process the

Table 1
Results From the Cull, Shaughnessy, and Zechmeister (1996), Cull (2000), and Carpenter and Delosh (2005) Studies

	Presentation schedule					
Experiment	1–5–9	5-5-5	0-0-0	1-3-5	3-3-3	0-0-0
Cull et al. (1996, Experiments 1–5)						
1. Test only	.33	.23	.17			
2. Test only	.27	.18	.11			
3. Test only	.34	.28	.28			
4. Test only	.72	.59	.59			
5. Test–study	.91	.91	.62			
Cull (2000, Experiment 1)						
Test only	.38	.34	.18			
Study only	.27	.29	.14			
Test-study	.48	.49	.19			
Carpenter & Delosh (2005, Experiment 2)						
Test only				.26	.30	.08
Study only				.12	.18	.08

information more fully, and hence this would lead to better retention. Because the longest retention interval in the expanded retrieval condition, compared with the equal-interval condition, should produce increased difficulty, with successful retrieval, this condition should produce the best long-term retention. Likewise, Landauer and Bjork (1978) originally argued that the benefit of expanded retrieval involves ensuring successful retrieval at longer and longer intervals, thereby increasing the difficulty of retrieval while still yielding success, that is, a type of shaping of retrieval success. One could also view these benefits within an encoding variability framework. Here, one might argue that with the passage of time and intervening information, there is an increased likelihood of encoding variability to play a role and hence a greater likelihood of distinct encodings. The notion here is that in a long-term retention test, stimuli that have more distinct encodings, afforded by the longest interval in the expanded condition, would especially produce a benefit compared with stimuli with less distinct encodings because of shorter spaced intervals (see, e.g., Glenberg, 1977).

It is important to note that both of the above perspectives suggest that there is an additional benefit beyond simply successful retrieval in the expanded condition compared with the equalinterval condition. However, to directly test this, one needs information about retrieval success during the original acquisition phase. Fortunately, Landauer and Bjork (1978) provided retrieval success data for each retrieval attempt during initial list presentation. (Cull et al., 1996, and Cull, 2000, did not report acquisition retrieval success rates, and Carpenter & DeLosh, 2005, reported retrieval success rates collapsed across attempts.) Figure 1 displays performance of an expanded condition compared with an equalinterval condition taken from Landauer and Bjork's first experiment. There is an obvious initial benefit of expanded retrieval that actually decreases across subsequent retrieval attempts. We find it interesting that the benefit of retrieval success in the expanded condition, compared with the equal-interval condition at the last attempt during list presentation, was actually larger than the effect found in final cued recall. Hence, instead of producing an increasing divergence of performance between the expanded and equalinterval conditions across the retrieval attempts, there appears to be some convergence. It would seem, on the basis of the above theoretical perspectives, that one should obtain an advantage of

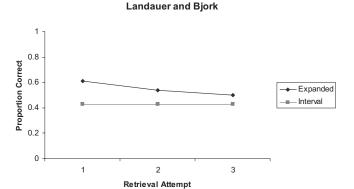


Figure 1. Results from Landauer and Bjork's (1978) study for the 1–5–10 expanding and the 5–5–5 equal-interval conditions.

expanding over equal-interval schedules above and beyond the retrieval success advantage produced during the acquisition phase. This does not appear to be the case, at least in the original Landauer and Bjork study. The expanded retrieval effect may be due to the fact that the first retrieval success occurs sooner after study, permitting greater retrieval success in the expanding relative to the equal-interval condition, as opposed to greater retrieval efficiency at the longest delays. We return to this issue in the General Discussion.

Aging and Alzheimer's Disease

Why might healthy older adults and individuals with DAT be a particularly important group to investigate the benefits of expanded retrieval? Of course, because of the clear episodic memory loss in these individuals, it is important to better understand the benefits of such expanded retrieval practice in these populations. In this light, we are surprised that there are no published studies that have directly compared expanded, equal, and massed retrieval schedules in a single experiment of healthy aging or Alzheimer's disease.

There has been some work investigating the benefits of spaced over massed presentation in healthy young and healthy older adults, as well as in individuals with early stage DAT. For example, in a study that factorially crossed retention interval and spacing, Balota, Duchek, and Paullin (1989) found that although older adults produced lower overall performance compared with younger adults, they produced the same effects of spacing and retention interval as younger adults. Balota et al. fit these data to Estes's (1955) stimulus sampling model (a precursor to notions of encoding variability mentioned above) and found that both the parameters that reflect the probability of storing contextual elements at a given point in time and the rate at which elements fluctuate across time (encoding variability) appear to change in healthy older adults, compared with younger adults. Because elements fluctuate at a slower rate in older adults, and this fluctuation has been tied to the spacing effect (see, e.g., Crowder, 1976, pages 294-297), it is possible that older adults may produce an increased benefit from the longest retention interval in an expanded retrieval condition, compared with the equal-interval condition.

Camp and colleagues (e.g., Abrahams & Camp, 1993; Camp, Foss, Stevens, & O'Hanlon, 1996; Camp & Stevens, 1990) have been particularly instrumental in exploring the benefits of expanded retrieval practice in individuals with DAT. These researchers have developed the SR (spaced-retrieval) training procedure that relies heavily on the benefits of expanded retrieval practice and corrective feedback. In this procedure, a patient may be given an item or event to remember and then, after a brief interval, may be given a retrieval test. If the individual does not recall the event, he or she is immediately provided feedback and required to give the correct answer, and the retention interval is decreased for the next trial. When a retrieval attempt is successful, then the retention interval is increased. Camp and colleagues have shown that this procedure is useful in learning names of objects (Abrahams & Camp, 1993), prospective memory tasks (e.g., Camp et al., 1996), and object-location pairs (Camp & Stevens, 1990) in individuals with DAT. There is no doubt that the results from the SR procedure have been striking in individuals with DAT in more applied settings.

Why might the SR procedure work so well in individuals with DAT? Camp et al. (2000) have argued that SR better taps a more implicit/automatic memory system than other mnemonic techniques. Compared with tasks that tap conscious recollection, implicit memory tasks (e.g., as reflected by repetition priming effects) appear to be relatively intact in both healthy older adults and individuals with early stage DAT (see, e.g., Balota & Duchek, 1989; Balota & Ferraro, 1996; Ferraro, Balota, & Conner, 1993; Gabrieli et al., 1999; Light & LaVoie, 1993). In this light, it is interesting to note that Spieler and Balota (1996) and Faust et al. (2001) have found that both healthy older adults and individuals with early stage DAT produce the same benefits of spacing and repetition in learning unrelated pairs of words in speeded naming performance. Because of the relatively minimal demand on attentional control systems, expanded retrieval seems ideally suited for individuals with relatively intact implicit memory systems but compromised attentional strategic systems (see Balota & Faust, 2001; Perry & Hodges, 1999).

Of course, to understand whether expanded retrieval affords an additional benefit over equal-interval condition, it would be useful to directly compare the two schedules in an applied setting. To our knowledge, Camp and colleagues (e.g., Camp et al., 2000) have not provided this comparison in these studies, but of course, this was not the goal of their research program. However, this issue was explored in a recent study by Hochhalter, Overmier, Gasper, Bakke, and Holub (2005) in a clinical setting. In their first experiment, 10 individuals diagnosed with DAT were presented pill names under five different schedules of retrieval practice: massed, uniform distributed (i.e., equal-interval), spaced/expanded with adjustment based on performance (similar to the Camp protocol described above), spaced/expanded without adjustment, and random. The results indicated that only 6 of 10 individuals showed long-term retention for at least one pill name, and there was no difference in the number of "learners" across the different schedules of retrieval practice. Thus the spaced/expanded retrieval condition did not show a benefit relative to any of the other schedules of practice. Although the Hochhalter et al. (2005) study is informative, the relatively small number of participants and more variable applied therapeutic setting indicates that further work is needed.

Overview of the Present Experiments

The present series of experiments compares massed, equal, and expanded retrieval schedules in relatively large samples of healthy young and older adults and in individuals with early stage DAT. Our focus is to further understand the benefits of expanded retrieval above and beyond the benefits of equal-interval spacing. Of course, one initial problem that arises in conducting studies with participant groups at varying levels of cognitive ability is the selection of an appropriate set of retrieval schedules. We needed to use a schedule that does not produce ceiling effects for our young adults and floor effects for our individuals with DAT. On the basis of pilot work, we implemented the schedules displayed in Table 2, which allowed a comparison of a 0-0-1-3-5 (expanded schedule) with a 0-0-3-3-3 (equal-interval schedule) and a 0-0-0-0-0(massed schedule). We included the initial two massed retrieval attempts to keep the individuals with DAT off of floor in the final cued-recall test. Given the relatively large differences in perfor-

Table 2

Examples of the Three Different Spacing Schedules

Massed	Equal interval	Expanded	
	GROUND-COLD	GROUND-COLD	
	GROUND-????	GROUND-????	
	GROUND-????	GROUND-????	
	•	•	
GROUND-COLD	•	GROUND-????	
GROUND-????	•	•	
GROUND-????	GROUND-????	•	
GROUND-????	•	•	
GROUND-????	•	GROUND-????	
GROUND-????	•	•	
	GROUND-????	•	
	•	•	
	•	•	
	•	•	
	GROUND-????	GROUND????	
	•	•	
	•	•	

mance across these groups of individuals, we anticipated that some of the participant groups may not benefit as much as other groups from the expanded retrieval schedule. However, we expected that the expanded schedule would produce better overall recall performance than the equal-interval schedule, which in turn would produce benefits compared with the massed condition. We also provide a measure of retrieval success during the acquisition phase to measure the benefit of expanded retrieval over the equal-interval condition during initial list presentation.

In the first experiment, we compared performance across three groups of individuals: young college-aged students, healthy older adults, and individuals with early stage DAT. We were initially interested in the benefits of expansion when no corrective feedback is given during list presentation, as in Landauer and Bjork's (1978) original first experiment. Our second and third experiments investigated the effects of the same variables but also provided corrective feedback during initial list presentation, as emphasized by Camp and colleagues (e.g., Camp, Bird, & Cherry, 2000).

Experiment 1

Method

Participants. A total of 103 individuals participated in this study. Seventy-four older participants were recruited from the Washington University Alzheimer's Disease Research Center (ADRC) for this study. All of the ADRC participants were originally screened for depression, hypertension, reversible dementias, and other disorders that could potentially produce cognitive impairment. The inclusionary and exclusionary criteria for DAT are consistent with the NINCDS-ADRDA criteria (McKhann et al., 1984). The severity of dementia was staged according to the Washington University Clinical Dementia Rating (CDR) scale (Berg, 1988; Hughes, Berg, Danziger, Coben, & Martin, 1982; Morris, 1993). According to this scale, CDR 0, 0.5, 1, 2, and 3 represent no dementia, very mild dementia, mild dementia, moderate dementia, and severe dementia, respectively. The CDR is based on a 90-min interview with both the participant and a collateral source. This interview both assesses the participant and relies on information available from the collateral source concerning the participant. This interview assesses the participants' cognitive abilities in the areas of memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care. Both the reliability of the CDR and the validation of the diagnosis (based upon autopsy) by the research team have been excellent (93% diagnostic accuracy) and well documented (e.g., Berg et al., 1998). Of the 74 participants, 31 were healthy, older controls (mean age = 75.8, SD = 8.78, range = 60–98; mean education = 14.9, SD = 3.18), and 43 individuals were classified as having very mild DAT (n = 34) or mild DAT (n = 9; mean age = 75.9, SD = 8.49, range = 51–91; mean education = 14.6, SD = 2.63). As shown in Table 3, the DAT individuals produced lower performance on all psychometric measures, with the exception of digit span forward, that were part of a standard psychometric battery that all the older individuals received at the ADRC. In addition to the 74 older participants recruited from the ADRC, we also recruited 29 young individuals from the undergraduate Psychology Department Subject Pool (mean age = 19.3, SD = 1.08, range = 18–22).

Apparatus. A PC was used to control the display of the stimuli and to collect participants' responses. We synchronized display of all stimuli with the vertical retrace of the monitor to control for presentation duration. The stimuli were displayed on a 14-in. video graphics array monitor.

Table 3

Experiment 1 Psychometric Means and Standard Deviations as a Function of Group

	Healthy control		D.	DAT	
Measure	M	SD	M	SD	
WMS					
Logical Memory	11.46	3.60	5.56	3.73*	
Forward Digit Span	6.47	1.11	6.19	1.14	
Backward Digit Span	5.17	1.37	4.26	1.36*	
Associate recall	16.05	3.71	9.94	4.06*	
Associate recognition	6.90	0.55	6.21	0.99*	
Word Fluency Test ^a	34.47	12.00	24.12	9.79*	
WAIS					
Information	22.43	4.72	15.60	6.40*	
Digit Symbol	50.87	11.00	34.33	14.10*	
Benton Copy Test—Form D ^b	9.83	0.59	9.05	1.53*	
Trail Making Test ^c					
Part A	34.07	9.58	61.05	32.90*	
Part B	91.43	40.50	133.80	47.90*	
Boston Naming Test ^d	55.80	5.60	45.50	13.00*	
AMNART	35.83	6.83	29.76	9.98*	
Animal Fluency ^e	21.27	6.65	12.05	6.22*	

Note. DAT = dementia of the Alzheimer's type; WMS = Wechsler Memory Scale (Wechsler & Stone, 1973); Associate recall = Paired Associate Learning Recall; Associate recognition = Paired Associate Learning Recognition; WAIS = Wechsler Adult Intelligence Scale (Wechsler, 1955); AMNART = American Version of the Nelson Adult Reading Test (Grober & Sliwinski, 1991).

^a Thurstone and Thurstone (1949).

^b Benton (1963).

^c Armitage

retrieval practices with one, three, and five intervening events (expanded: 0-0-1-3-5). As noted, we included the two massed retrieval practices at the beginning of each sequence to increase the likelihood that the individuals with DAT correctly encoded the word pair at study.

We counterbalanced critical items across lists to ensure each word pair occurred in each of the three study list conditions. Presentations and retrieval practices of critical word pairs were interwoven throughout the list to maintain the spacing of the study list conditions. In addition, we embedded filler items throughout the study list to ensure the required spacing of items. Each list was composed of 80 events (study pairs, retrieval practices, and filler items). Finally, within the study list structure, a particular study list condition (e.g., massed, interval, expanded) was never presented twice in a row, and on average massed, interval, and expanded items were equally spaced throughout the study list (e.g., all of the massed items did not occur at the beginning or end of the list; average list position for massed = 38.3, interval = 42.3, expanded = 40.5).

Procedure. All participants were tested individually. They were told that a list of items would be presented for study and testing and that this would be followed by a later recall test. For each study presentation, the word pair was presented in the middle of the screen for 3 s. For each retrieval practice, the first word was presented with a blank (e.g., GROUND—_____) until the participant made a response or 15 s elapsed. Participants were instructed to verbally respond with the corresponding target word or to guess if they were uncertain. Correct or incorrect feedback was not provided during retrieval practices during the acquisition phase.

Following the acquisition phase, a 5-min break was given, during which participants engaged in a conversation. During this time the experimenter asked the participants open-ended questions, such as "What is your favorite vacation spot?" and "What is your favorite food?" A cued-recall test was then administered. The first word for each of the critical word pairs was presented (e.g., GROUND—_____) until the participant made a verbal response or 15 s elapsed. Participants were instructed to try to retrieve the target word that was paired with the cue word during study.

Results

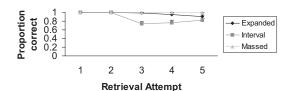
Proportion correct retrieval during the acquisition phase. The proportion of correct retrieval during the acquisition phase, as a function of group and spacing condition, is displayed in Figure 2. As shown in Figure 2, as expected, the massed condition produced consistently high performance across retrieval attempts during study for all groups. Moreover, performance for all groups was equivalent on the equal-interval and the expanded condition on the first two retrieval attempts (which were identical), but then there was a large difference, such that the expanded retrieval condition produced a benefit, compared with the equal-interval condition, with this difference decreasing somewhat across the subsequent two retrieval events. These results are overall consistent with the Landauer and Bjork (1978) observations. Also, as shown here, there is a clear benefit of expanded retrieval over equal interval at the last retrieval event across all groups. This advantage of the expanded condition (16% collapsed across all groups) is actually larger than the benefit on the last retrieval attempt found in the original Landauer and Bjork study (7%, on the basis of interpolation from their Figure 2).

The results from a 3 (group) \times 3 (spacing) \times 5 (retrieval attempt) mixed-factor analysis of variance (ANOVA) supported the above observations. As expected, there were highly reliable main effects of group, F(2, 100) = 24.54, MSE = .251, p < .001; spacing, F(2, 200) = 81.84, MSE = .086, p < .001; and retrieval attempt, F(4, 400) = 143.78, MSE = .022, p < .001. However,

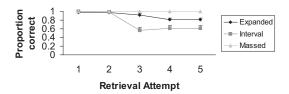
^a Thurstone and Thurstone (1949). ^b Benton (1963). ^c Armitage (1946). ^d Goodglass and Kaplan (1983b). ^e Goodglass and Kaplan (1983a).

p < .01

Young Retrieval Attempts No Feedback



Healthy Controls Retrieval Attempts No Feedback



DAT Retrieval Attempts No Feedback

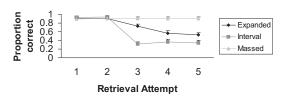


Figure 2. Mean proportion correct during the acquisition phase for the young, older healthy control, and dementia of the Alzheimer's type (DAT) individuals, as a function of spacing, and retrieval attempt.

these main effects were qualified by a Group X Spacing X Attempt interaction, F(16, 800) = 6.29, MSE = .018, p < .001, which as shown in Figure 2, indicates that there was a larger effect of spacing for the DAT individuals, especially at retrieval attempts 3-5, compared with the healthy young and older adults. This three-way interaction was pursued by a series of ANOVAs comparing the spacing effect as a function of group across retrieval attempts 3, 4, and 5, because all groups were at ceiling for retrieval attempts 1 and 2. First, comparing young versus healthy older controls, there was a significant Group \times Spacing interaction, F(2,116) = 4.31, MSE = .367, p < .02, which, as shown in Figure 2, reflected the finding that there was a larger difference between massed and the two spaced conditions in the healthy older adults compared with the younger adults. In addition, comparing healthy older controls versus the DAT group, there also was a reliable Group \times Spacing interaction, F(2, 144) = 3.13, MSE = .449, p <.05, which indicated that there was a larger difference between massed and the two spaced conditions in the DAT individuals, compared with the healthy older adults. A separate ANOVA that considered only the expanded and equal-interval conditions indicated that there was not a reliable Group X Spacing interaction either in the ANOVA comparing young versus healthy older adults F(1, 58) = 1.11, MSE = .114, p = .30, or in the ANOVAcomparing healthy older adults versus DAT individuals, F(1,72) = 0.05, MSE = .183, p = .83.

Proportion correct final recall. The proportion correct in the final cued-recall test as a function of spacing is displayed in Figure 3. The results here are quite striking. In particular, there is a large benefit of the two spaced conditions, compared with the massed condition, with no hint of a difference between the expanded and equal-interval conditions for any group of participants. The results of a 3 (group) × 3 (spacing condition) mixed-factor ANOVA yielded significant main effects of group, F(2, 100) = 37.58, MSE = .158, p < .001, and spacing, F(2, 200) = 18.88, MSE =.056, p < .001. The Group \times Spacing interaction did not reach significance, F(4, 200) = 1.24, MSE = .056, p < .29. It is important to note that post hoc t tests comparing spacing conditions for each of the three groups separately indicated higher cued recall performance comparing the expanded versus massed condition for the young, t(28) = 2.75, p < .01; old, t(30) = 3.61, p < .01.001; and DAT individuals, t(42) = 2.68, p = .011, respectively, and comparing the interval versus massed condition for the young, t(28) = 1.97, p = .059; old, t(30) = 3.58, p < .001; and DAT individuals, t(42) = 2.99, p < .005, respectively. However, there was no difference in final cued recall performance comparing the expanded versus equal-interval conditions for the young, older control, or DAT participants (ps > .30).

Discussion

The findings from Experiment 1 are quite clear. First, the results from the initial acquisition phase indicated that indeed the expanded retrieval condition produced higher retrieval success than the equal-interval condition. As noted, performance on the last retrieval attempt in the expanded condition was considerably higher (16%) than the equal-interval condition. However, instead of maintaining the advantage of expanded over equal-interval spacing that was obtained during list presentation, this advantage was lost in the final cued-recall test. Of course, it is also clear that spacing of retrieval produced a large effect on later recall performance, that is, both spaced conditions showed a large advantage, compared with the massed condition, across all three groups. The results from Experiment 1 are especially intriguing in light of the fact that a very similar pattern was found in final recall performance across three groups of participants that produced dramatic differences in overall recall performance.

Experiment 2

In the second experiment, we were interested in the role of corrective feedback on the influence of study schedules. In this

Correct Recall at Test No Feedback

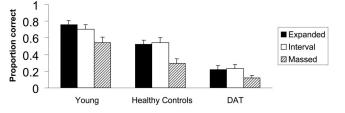


Figure 3. Mean final cued recall performance as a function of group and spacing schedule. DAT = dementia of the Alzheimer's type.

experiment, if the participant's retrieval attempt for the target word failed when the participant was presented with the cue during the acquisition phase, the experimenter provided the correct response. We were especially interested in the possibility that the feedback would bring participants in the equal-interval condition up to the level of the expanded condition during the acquisition phase. Moreover, it is possible that the feedback may diminish some of the large group differences found in Experiment 1 during the acquisition phase. The role of corrective feedback is particularly important because, as noted earlier, Camp et al. (2000) suggested that their SR method is dependent upon feedback during the initial list presentation. It may be the case that feedback would be especially important for the DAT individuals.

Method

Participants. A total of 109 individuals participated in this study. Seventy-two of the participants were recruited from the Washington University ADRC for this study. Of the 72 participants, 38 were healthy older controls (mean age = 79.2, SD = 7.63, range = 63–95; mean education = 13.4, SD = 3.09), and 34 were diagnosed with very mild (n = 26) or mild DAT (n = 8; mean age = 78.4, range = 68–98, SD = 7.46, mean education = 14.3 years, SD = 3.05). The psychometric performance for the healthy older adults and the individuals with DAT are presented in Table 4. As shown here, the two groups are different on all psychometric tests, with the exception of forward digit span and associate recognition tests.

In addition, 37 young undergraduate students were recruited from the Washington University Psychology Department Subject Pool to participate in this study (mean age = 20.7, SD = 3.36, range = 18-37).

Table 4
Experiment 2 Psychometric Means and Standard Deviations as a Function of Group

	Healthy	control	DAT	
Measure	M	SD	M	SD
WMS				
Logical Memory	9.86	3.77	5.56	4.08*
Digits Forward	6.24	1.22	5.97	1.34
Digits Backward	5.08	1.22	4.03	1.27*
Associate Recall	14.68	3.67	11.26	3.72*
Associate Recognition	6.92	0.36	6.67	0.96
Word Fluency Testa	28.97	10.40	24.12	8.86*
WAIS				
Information	21.53	4.11	16.91	5.33*
Digit Symbol	47.32	10.60	36.52	13.90*
Benton Copy Test—Form D ^b	9.89	0.31	9.18	1.55*
Trail Making Test ^c				
Part A	37.44	13.20	60.41	36.8*
Part B	95.63	43.90	123.00	43.50*
Boston Naming Test ^d	55.47	4.26	49.06	9.73*
NART	35.39	7.42	28.03	11.00*
Animal Fluency ^e	19.26	6.02	14.38	5.94*

Note. DAT = dementia of the Alzheimer's type; WMS = Wechsler Memory Scale (Wechsler & Stone, 1973); Associate recall = Paired Associate Learning Recall; Associate recognition = Paired Associate Learning Recognition; WAIS = Wechsler Adult Intelligence Scale (Wechsler, 1955); AMNART = American Version of the Nelson Adult Reading Test (Grober & Sliwinski, 1991).

^a Thurstone and Thurstone (1949).

^b Benton (1963).

^c Armitage

^a Thurstone and Thurstone (1949). ^b Benton (1963). ^c Armitage (1946). ^d Goodglass and Kaplan (1983b). ^eGoodglass and Kaplan (1983a).

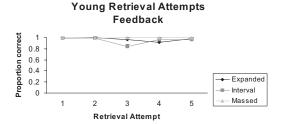
Materials and procedure. All materials and procedures were identical to Experiment 1, with one exception. In Experiment 2, participants were given corrective feedback during the acquisition phase. Participants were provided with the correct response if they produced the incorrect response, indicated that they did not remember the correct response, or the 15-s retrieval period ended with no response. When participants were given corrective feedback, they were not asked to repeat the correct response aloud.

Results

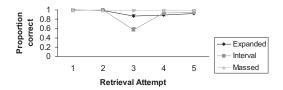
Proportion correct recall during the acquisition phase. The proportion of correct retrieval during the acquisition phase as a function of group, spacing, and retrieval attempt is displayed in Figure 4. Overall, by comparing Figure 4 and Figure 2, one can see that the feedback did increase successful retrieval during the acquisition phase, especially for the last couple of retrieval attempts across all groups, wherein feedback could play a role. In fact, if one compares Experiments 1 versus 2 for retrieval attempts 4 and 5, there is a main effect for Experiment, F(1, 206) = 51.59, MSE = 6.07, p < .001. In contrast to Experiment 1, as shown in Figure 4, there appears to be very little difference in performance on the fourth and fifth retrieval attempts between the equal-interval and expanded retrieval conditions, where it is apparent that the young and older adults are near ceiling in both of these conditions, which was reflected by an Experiment \times Spacing interaction, F(2,412) = 42.12, MSE = 2.42, p < .001, on the fourth and fifth retrieval events during the acquisition phase. Any difference after the third retrieval attempt in the expanded and the equal-interval condition is drastically reduced in the next couple of retrieval attempts in this experiment, due to feedback. Indeed, as shown in Figure 4, there is a particularly dramatic influence of feedback for the equal-interval condition after the third retrieval attempt. This effect becomes larger in the healthy older adults, compared with young adults, and becomes even more exaggerated in the DAT individuals.

The results of a 3 (group) \times 3 (spacing) \times 5 (retrieval attempt) mixed factor ANOVA yielded significant main effects of group, F(2, 106) = 34.59, MSE = .074, p < .001; spacing, F(2, 212) =79.17, MSE = .028, p < .001; and retrieval attempt, F(4, 424) =77.59, MSE = .020, p < .001. These main effects were qualified by a reliable three-way interaction among group, spacing, and retrieval attempt, F(16, 848) = 7.10, MSE = .017, p < .001. As shown in Figure 4, this interaction reflected the finding that the performance in the equal-interval condition decreased during the third retrieval attempt, and this decrease was greater in the healthy old compared with the healthy young and greater in the DAT individuals compared with the healthy older adults. This three-way interaction was pursued by a series of ANOVAs comparing the spacing effect as a function of group at retrieval attempt 3, where feedback has the first opportunity to influence performance, and the inferences are not limited by ceiling effects. First, comparing young versus healthy controls, there was a significant Group X Spacing interaction, F(2, 146) = 11.65, MSE = .321, p < .001, which, as shown in Figure 4, reflected the finding that there was a larger difference in the decrease in the equal-interval condition at retrieval attempt 3 for the healthy older adults compared with the younger adults. In addition, comparing healthy controls versus the DAT group, there also was a reliable Group \times Spacing interaction, F(2, 140) = 3.80, MSE = .196, p < .03, which indicated that there

p < .04



Healthy Controls Retrieval Attempts Feedback



DAT Retrieval Attempts Feedback

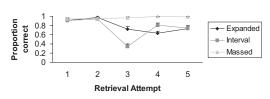


Figure 4. Mean proportion correct during the acquisition phase for the young, older healthy control, and dementia of the Alzheimer's type (DAT) individuals, as a function of spacing, and retrieval attempt.

was a larger difference in the decrease in the equal-interval condition at retrieval attempt 3 for the DAT individuals compared with the healthy older adults.

Proportion correct final recall. The proportion correct recall at the final cued-recall test is displayed in Figure 5. As shown, there is a large benefit of the expanded and the equal-interval condition over the massed condition, that is, a spacing effect, with the expanded and equal-interval conditions yielding an identical pattern. This pattern is very similar to the results obtained in Experiment 1. The results of a 3 (group) \times 3 (spacing condition) mixed-factor ANOVA yielded significant main effects of group, F(2, 106) = 48.41, MSE = .138, p < .001, and spacing, F(2, 106) = .001212) = 57.09, MSE = .041, p < .001, but no hint of an interaction between group and spacing, F(4, 212) < 1.00. Post hoc t tests indicated higher cued recall performance in the expanded compared with the massed condition for the young, t(36) = 6.48, p <.001; old, t(37) = 5.05, p < .001; and DAT individuals, t(33) =4.41, p < .001, and higher recall in the equal-interval compared with the massed condition for the young, t(36) = 7.73, p < .001; old, t(37) = 4.52, p < .001; and DAT individuals, t(33) = 5.42, p < .001. However, there was again no reliable difference in final cued-recall performance comparing the expanded and the equalinterval conditions for any of the groups (ps > .38).

Discussion

The results from the second experiment are noteworthy along a number of dimensions. First, the corrective feedback provided during the acquisition phase was successful in increasing accurate retrieval, especially in the equal-interval condition. As shown in Figure 4, the effect of feedback was most evident after the first spaced retrieval in the equal-interval condition, that is, after the third retrieval event. This finding appears largest for the DAT individuals. Of course, the considerable drop in performance at the third retrieval event is reasonable, as in the equal-interval condition, this is the first retrieval event after three intervening items. More interesting is the powerful influence of corrective feedback that pushes performance in the equal-interval condition slightly above the expanded retrieval condition at the fourth retrieval attempt. Second, the benefit of feedback produced a clear influence on final cued recall performance, relative to the results of Experiment 1, especially for the spaced conditions. In particular, in the expanded condition, there is approximately a 13% influence of feedback across Experiments 1 and 2 (p = .007) and a 15% influence of feedback for the equal-interval condition across Experiments 1 and 2 (p = .002). Third, and most important, there was no hint of a difference between the expanded and equalinterval conditions for any group of participants, even though final recall performance varied by as much as 50% between the healthy young individuals and the early stage DAT individuals.

Experiment 3

The lack of difference between the expanded and equal-interval conditions in cued-recall performance in both Experiments 1 and 2 is somewhat surprising but is consistent with more recent work on this topic reviewed in the Introduction. It is possible that the inclusion of the two immediate retrieval attempts may have masked a benefit of expanded retrieval over the equal-interval condition. Of course, we initially included these immediate retrieval attempts to ensure that all participant groups (especially the DAT individuals) were able to successfully encode the pairs. To address the potential influence of the initial retrieval attempts, in the third experiment, we eliminated the two initial massed retrieval attempts. Hence, instead of 0-0-1-3-5, 0-0-3-3-3, and 0-0-0-0-0 conditions, we included 1-3-5, 3-3-3, and 0-0-0 conditions. Again, participants were provided corrective feedback if they were unable to retrieve the response item during a retrieval

Correct Recall at Test Feedback

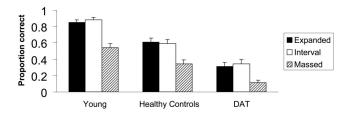


Figure 5. Mean final cued recall performance as a function of group and spacing schedule. DAT = dementia of the Alzheimer's type.

attempt. In this study, we targeted only a group of healthy younger adults and a group of healthy older adults.

Methods

Participants. A total of 25 individuals participated in this study. Fifteen participants were recruited from the Washington University Psychology Department Older Adult Subject Pool for this study (mean age = 77.4, SD = 4.69, range = 70-86), and 10 young undergraduate students were recruited from the Washington University Psychology Department Subject Pool (mean age = 20.4, SD = 3.24, range = 18-29).

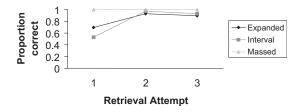
Materials and procedure. All study list materials (word pairs) and the cued-recall tests were identical to Experiment 2, with the exception of the two initial massed retrieval attempts. Hence, each study list consisted of 12 critical items that occurred under one of three spacing schedules (massed, interval, expanded). As noted, the three spacing schedules included the following: massed (0-0-0), interval (3-3-3), and expanded (1-3-5). Presentations and retrieval practices of critical word pairs were interwoven throughout the list. To maintain the spacing of the conditions filler items were embedded throughout the study list. This resulted in a slightly shorter study list of 70 events (study pairs, retrieval practices, and filler items).

All study and testing procedures were identical to Experiment 2. Corrective feedback was provided during the retrieval attempts in the acquisition phase.

Results

Proportion correct recall during the acquisition phase. The proportion of correct recall as a function of group, spacing, and retrieval attempt during the initial acquisition phase is presented in Figure 6. Overall, it appears that there was relatively little influence on the pattern of spacing effects in this experiment compared with Experiment 2. This can be seen best by comparison of the third, fourth, and fifth retrieval attempts in Figure 4 (i.e., excluding the two initial retrieval attempts) with the first, second, and third retrieval attempts in Figure 6. Specifically, there is again a dramatic influence of feedback, which is much larger for the equal-

Young - Retrieval Attempts



Older - Retrieval Attempts

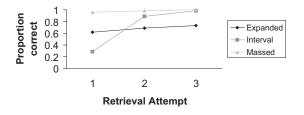


Figure 6. Mean proportion correct during the acquisition phase for the young and older individuals, as a function of spacing, and retrieval attempt.

interval condition than in the expanded condition, especially for the older adults.

The results of a 2 (group) \times 3 (spacing) \times 3 (retrieval attempt) mixed factor ANOVA yielded main effects of spacing, F(2, 46) = 18.87, MSE = .064, p < .001, and retrieval attempt, F(2, 46) = 65.58, MSE = .020, p < .001, with the main effect of group approaching significance, F(1, 23) = 3.95, MSE = .118, p = .059. These main effects were qualified by two interactions. First, there was a Spacing \times Retrieval Attempt interaction, F(4, 92) = 17.81, MSE = .032, p < .001, which indicated that the difference between the spaced schedules and the massed schedules decreased across retrieval attempts. Second, the Group \times Spacing \times Retrieval Attempt interaction approached significance, F(4, 92) = 2.10, MSE = .032, p = .087, which indicated that performance as a function of retrieval attempt and spacing was modulated more in the older adults than in the young adults. This was again similar to the pattern observed in Experiment 2.

Proportion correct final recall. The proportion correct recall at the final cued-recall test is displayed in Figure 7. As shown in Figure 7, the results are again quite clear. There was a large effect of spacing but little difference between the expanded and equalinterval conditions, with some tendency for the equal-interval condition to produce slightly higher performance than the expanded condition.

The results of a 2 (group) \times 3 (spacing condition) mixed-factor ANOVA yielded main effects of group, F(1, 23) = 17.26, MSE = .089, p < .001, and spacing, F(2, 46) = 13.57, MSE = .066, p < .001. Again, the Group \times Spacing interaction did not approach significance (F < 1.00). Post hoc t tests comparing the two spaced conditions with the massed condition for each group indicated higher cued-recall performance in the expanded versus massed condition (ps < .02) and higher cued-recall performance in the equal-interval condition was slightly higher than the expanded condition, this difference did not reach significance for either the younger or older adults (ps > .40).

Discussion

The results from Experiment 3 are quite clear and nicely replicate comparable conditions in Experiment 2. Specifically, during the initial acquisition phase, there was a substantial benefit of feedback in the equal-interval condition after the initial spaced presentation, whereas there was relatively little benefit in the expanded retrieval condition. Turning to the final cued-recall test, there was a substantial benefit of the two spaced conditions over the massed presentation condition. Again, there was no difference between the expanded and the equal-interval condition, and if anything, there was a slight benefit for the equal-interval condition. Hence, it does not appear that the presence of the initial massed retrieval events in the schedules used in the first two experiments limits the implications of these results.

General Discussion

The results from the present series of experiments are quite clear: First, Experiment 1 yielded a large effect of spacing schedule, with the expanded retrieval condition yielding higher performance than the equal-interval condition during the initial acquisi-

Correct Recall at Test

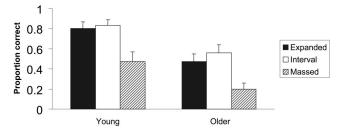


Figure 7. Mean final cued recall performance as a function of group and spacing schedule.

tion phase. In fact, the benefit of expanded retrieval over equalinterval during the acquisition phase was larger in this experiment than in the original Landauer and Bjork (1978) study. However, this benefit did not carry over into final cued-recall performance. Specifically, there was no hint of a difference between the expanded retrieval condition and the equal-interval condition in the final cued-recall test, even though both conditions produced large benefits over the massed condition. Second, in Experiment 2, we provided participants with corrective feedback after each trial in which the correct response was not produced. This was in part motivated by the possibility that feedback may have a particularly powerful benefit under expanded retrieval conditions. The results indicated that feedback produced a considerable benefit that was primarily localized for the equal-interval condition (relative to the expanded retrieval condition) after the initial spaced retrieval event. Although there was also a clear benefit of feedback on later cued recall, that is, overall higher performance in Experiment 2 than Experiment 1, feedback did not produce a difference between the expanded and equal-interval conditions, with both conditions again producing large benefits over the massed condition. Finally, we conducted Experiment 3 to ensure that the presentation of the two initial massed retrieval attempts in the first two experiments did not compromise the results. The results again produced the same pattern, that is, a large spacing effect, but no difference between the expanded and the equal-interval conditions.

Possibly, the lack of a difference between the expanded and equal-interval conditions in Experiment 1, may have been due to the specific schedules employed that minimized the benefits of expanded retrieval. For example, as noted, because of the large variation in memory performance across our groups, we included two initial massed retrieval attempts to ensure the items were initially encoded. However, this did not appear to change the influence of expansion during the acquisition phase. Specifically, as shown in Figure 2, it is quite clear that the expanded condition did indeed produce the desired higher performance than the equalinterval condition during the initial acquisition phase. In fact, by the last retrieval attempt, the expanded condition produced benefits in absolute terms of 19% for both the healthy older adults, t(30) =3.06, p = .005, and the DAT individuals, t(42) = 3.41, p = .001, and 9% for the young adults, t(28) = 2.13, p = .042. Hence, compared to the equal-interval condition, expanded retrieval indeed accomplished the goal of maintaining high levels of retrieval success during the acquisition phase at a longer retention interval (5 vs. 3 intervening items at the last retrieval event). However, there was no hint of a difference between the expanded and equal-interval condition in the final cued-recall test.

It is also possible that the expanded retrieval advantage was not obtained because participants were simply at an insensitive area of the measurement scale (i.e., possibly masked by ceiling or floor effects). In retrospect, one advantage of the present experiments is that each study included a large group of individuals at widely varying levels of overall performance during the initial acquisition phase and the later cued-recall test. Therefore, it is unlikely that the present failure to obtain an advantage of the expanded condition over the equal-interval condition for any group of participants was due to participants being at performance levels that somehow masked the effect.

There are two further points to note: First, Logan (2004) has recently explored a series of expanded and equal-interval conditions in healthy young and older adults. For example, she compared schedules of 1-2-3 versus 2-2-2, 1-3-5 versus 3-3-3, and 4-4-4 versus 1-3-8. Logan found no evidence of a benefit of the expanded condition over the equal-interval condition extending into a final recall condition across these schedules, and in fact found some evidence of greater forgetting in the expanded retrieval condition at a one day retention interval. As noted earlier, Cull (2000) also found no difference between the 5–5–5 and 1–5–9 schedules. Hence, it does not appear that the present lack of an effect is due to the particular schedules used in the present study. Second, Carpenter and DeLosh (2005) used face-name pairs as the stimulus items, and as noted above provided some evidence that the equal-interval condition was slightly better than the expanded condition; hence, the present lack of a benefit of the expanded retrieval condition extends beyond low-associate word pairs (also see Cull, 2000).

¹ One might be concerned that the type of intervening event (i.e., the presentation of a new target pair vs. a retrieval attempt) within the study list differed between the interval and expanded spacing conditions, thus differentially influencing the timing of the acquisition phase before recall. However, this did not appear to be the case. The percentage of intervening events between the presentation of the target pair and the recall test that involved the presentation of new items (i.e., new target pairs or filler pairs), was 46% in the interval condition and 40% in the expanded condition in Experiments 1 and 2. The percentage of intervening events that involved retrieval attempts was 54% in the interval condition and 60% in the expanded condition in Experiments 1 and 2. The percentages of new items and retrieval attempts as intervening events in the equal interval condition (59.5% and 40.5%, respectively) were identical to those in the expanded condition (59.5% and 40.5%, respectively) in Experiment 3. It is also important to note that even though participants were given 15 s to make a response during a retrieval attempt, participants' response latencies for successful trials were relatively fast and did not vary across conditions, as one might expect if the benefits of spacing reflect a more automatic implicit process. Specifically, very few participants across all the groups had correct response latencies over 4 s (0.6%, 1.4%, or 2.2% of participants in the expanded, interval, and massed conditions, respectively). Finally, and most important, the expanded retrieval condition produced marked benefits over the equal-interval condition in the first experiment study during the acquisition phase. Hence, any variability from intervening events in the present study clearly did not disrupt the expected maintenance of high retrieval success that the expanded retrieval condition affords during acquisition.

Turning to the role of feedback, we considered the possibility that such corrective information would be especially beneficial for healthy older adults and individuals with DAT. To our knowledge, the present experiments are the first to directly investigate the advantage of feedback in the expanded and a comparable equalinterval condition. Although there was a clear influence of feedback that differed among groups during the acquisition phase, these differences did not translate into a long-term benefit of the expanded retrieval condition compared with the equal-interval condition as reflected on the final cued-recall test. Again, there was a large spacing effect, but no difference between the expanded and equal-interval conditions. Thus, there does not appear to be evidence of a heightened benefit of expanded retrieval when feedback is provided, at least within the present schedule constraints. It is important to note here that although corrective feedback was provided in our experiments, participants were not required to repeat the correct answer at the time of feedback, and the schedule was not adjusted as a function of accuracy, both of which are changes from the typical protocol used by Camp and colleagues (e.g., Camp et al., 2000).

Theoretical Implications

No feedback conditions. One of the appealing aspects of the expanded retrieval procedure is that it helps maintain a high level of retrieval success by gradually increasing the interval between retrieval attempts. Hence, by the last retrieval attempt, one is accurately recalling the item at a greater retention interval and at a higher accuracy level, compared with the equal-interval condition (i.e., five intervening items compared with three intervening items in the present study). Because the individual is better retrieving information at a greater retention interval in the expanded condition compared with the equal-interval condition at the last retrieval attempt, one might expect this effect to become larger in a final cued-recall test. The notion is that the strength of the underlying memory traces at the last retrieval event must be greater in the expanded compared to the equal-interval condition, because the items were accurately retrieved at a longer retention interval in the former. However, this advantage was lost in the final recall test. If one takes into consideration the higher retrieval success rate during acquisition in the expanded condition, and the equal-performance on the final recall test, one would argue that there is greater forgetting in the expanded condition, compared to the equalinterval condition. As noted earlier, even in the original Landauer and Bjork (1978) study, there was little evidence of an increased benefit of the expanded retrieval condition compared with the equal-interval condition in the final cued-recall test, when one takes into consideration the initial difference between these two conditions during the acquisition phase. We are unaware of any evidence that shows the expected boost in final recall of an expanded retrieval condition over an equal-interval condition above and beyond the initial influence on retrieval success. In fact, the relatively high retrieval success during the acquisition phase may be attributable to the initial first short retention interval in the expanded condition (i.e., with only one intervening item). Ultimately, it is possible that participant's long-term memory may actually benefit more from the initial spacing of three intervening items in the equal-interval condition, compared with the relatively short spacing of one intervening item in the initial expanded condition. In this sense, the equal-interval condition may represent a condition in which there is a long-term benefit of three spaced retrieval attempts (3–3–3), whereas, the expanded retrieval condition only represents two meaningful spaced retrieval attempts (3 and 5). Hence, the short-term gain in performance for the initial retrieval event in the expanded condition may be compensated by a long-term loss.

Although there are a number of viable theoretical accounts of the spacing effect (see Cepeda, Pashler, Vul, Wixted, & Rohrer, in press, for a review; Verkoeijen, Rikers, & Schmidt, 2004), the potential loss of the expanded retrieval advantage observed during acquisition in a long-term recall test could be viewed within an encoding variability framework. The encoding variability account suggests that contextual elements fluctuate across time and that the likelihood of recalling any single item is dependent upon how much overlap there is in the contextual information available at the time of encoding and at the time of retrieval. For example, one pattern of results that has been viewed as strongly supportive of encoding variability accounts is the Spacing × Retention Interval interaction. Specifically, at short retention intervals massed practice actually produces better performance than spaced practice, whereas, at longer retention intervals one finds the opposite pattern (see, e.g., Balota et al., 1989; Glenberg, 1977). The encoding variability account handles this interaction quite nicely since immediate tests will produce large benefits of the double encoding of an event because the contextual elements have had relatively little time to fluctuate, and so there will be considerable overlap between the retrieval context and the encoded context. On the other hand, on a delayed memory test, the contextual elements during retrival are going to be relatively new, and so the more distinct contextual elements encoded, as with spaced practice, the better the delayed memory performance.

Consider the encoding variability account of the present results. Since the initial retrieval event in the expanded condition involves only one intervening item, whereas in the equal-interval condition, this retrieval event involved three intervening items, it is likely that there is going to be an increasing influence of contextual fluctuation in the equal-interval condition. It is possible that the initial retrieval event in the expanded condition is having an influence more like a massed retrieval event, with relatively little opportunity for contextual fluctuation. This will of course help maintain performance on an immediate test (just like massed spacing benefits immediate testing) but may not produce a long-term benefit (just like massed spacing impairs delayed testing). Of course, one would also have to argue here that there is relatively little difference between the contextual variability of five intervening items in the expanded retrieval condition and the three intervening items in the equal-interval condition. If this account were correct, then one might expect that testing at greater retention intervals might actually reverse the short-term benefits produced by the expanded retrieval condition during the acquisition phase. As noted above, Carpenter and DeLosh (2005) found evidence of a reliable advantage of an equal-interval condition over an expanded condition at a 5-min filled retention interval, even though the expanded condition afforded a 20% benefit (collapsed across retrieval events) over the equal interval during the acquisition phase. Likewise, Logan (2004) found at a 1-day retention interval, the equal-interval condition produced better performance than expanded retrieval in young adults. Finally, Karpicke and Roediger (2005) have recently presented evidence that expanded retrieval (1–5–9) produces a benefit over equal-interval (5–5–5) condition at a 10-min delay, but the effect reliably reverses at a 2-day interval. Clearly, further work is needed to test the influence of different retention intervals across different equal and expanded schedules. However, the potential tradeoff in short-term benefits versus long-term losses across these schedules is reminiscent of the spacing by retention interval interaction, which ultimately may be best accommodated by encoding variability theory.

Corrective feedback conditions. Another intriguing aspect of the present results was the remarkably large boost in performance during the acquisition phase because of corrective feedback provided after the third retrieval attempt in the equal-interval condition (Figures 4). This effect occurred for all groups but was most exaggerated in the DAT individuals, where there was a 47% increase in retrieval accuracy between the third and fourth retrieval attempts because of the feedback. There was relatively little evidence of a benefit of feedback after the third retrieval attempt in the expanded condition, even though the DAT individuals were clearly not at ceiling performance. This pattern appears to suggest that the utility of feedback may be dependent on the availability of the recently retrieved information (Jacoby, 1978). In the expanded condition after the third retrieval attempt, the retention interval was only one intervening item, and so the item was in a relatively active state, that is, little time for contextual fluctuation. It is possible that participants did not fully reengage retrieval, because the corrective feedback produced a relatively strong match between the item recently presented and the feedback, thereby producing a strong familiarity signal. On the other hand, in the equal-interval condition, after three intervening items, the item may no longer be in an available state, and so it is possible that the corrective feedback was more likely to engage attention demanding retrieval operations. The notion here is that it is possible that the familiarity of the feedback information affords a cue to the attentional system to engage in more elaborative processing, that is, within the encoding variability framework encoding more of the contextual elements. Of course, future work is needed to verify this balance of retrieval difficulty and the utility of feedback.

Applied Implications

The failure to obtain a difference between the expanded and the equal-interval retrieval conditions would appear to bring into question the utility of expanded retrieval in more applied settings. However, we believe that such a conclusion is too strong for the following reasons: First, although it is unclear whether there is anything special about expanded retrieval above and beyond spacing at equal-intervals for long-term retention, expanded retrieval may be a useful way to implement spacing. That is, it may be easier to implement a schedule with increasing lags than a priori assuming one can pick a useful equal-interval schedule. Second, we believe the success that participants receive during the acquisition phase in the expanded retrieval schedule may be a particularly important reinforcing event and hence increase the motivation to stick with an expanded retrieval schedule, especially in memory-impaired populations where retrieval success may be limited. This may be the way to reap at least most of the benefits of spaced practice. Third, it is still quite possible that comparing schedules that have more than three attempts would produce better performance in the expanded condition. For example, one might expect some benefit of a 1-3-5-7-10-14 schedule compared with a 5-5-5-5-5-5 schedule. Of course, implementing such schedules may be difficult in a well-controlled manner, but clearly this is important since acquisition of information often extends beyond three learning events. Finally, one must also be cautious in extending the present research to therapeutic settings, with memory-impaired individuals. For example, the experimental protocol used in the present studies differs in various ways from the Camp (2000) protocol. Multiple targets were presented during acquisition, rather than a single target, and Camp and colleagues used a shaping procedure wherein the retrieval practice expands and contracts depending on the success of retrieval at each trial. Ultimately, this may be the best procedure to implement spaced retrieval in memory-impaired individuals (but see Hochhalter et al., 2005). Clearly, further work is necessary to isolate the specific characteristics of spaced schedules that are beneficial with memory-impaired individuals.

In summary, across all experiments, the largest effect obtained was between the two spaced conditions and the massed condition. The consistency of the size of this effect across groups, different spacing conditions, and feedback conditions converges on the fundamental importance of spacing in learning. For example, there were also large differences in accuracy in the two spaced conditions during the acquisition phase and yet no appreciable difference in long-term retention. This suggests another dissociation between the short-term gains of spacing and the more long-term benefits. Also, the finding that there was no evidence of a Group \times Spacing interaction, where the participants ranged from healthy young adults to individuals with early stage DAT, suggests that the spacing effect is relatively immune to global cognitive decline or baseline episodic memory performance. Finally, it is unlikely that the present results are due to strategies that have developed across list presentations, because one would have expected group differences in such strategies. In this light, the present results are consistent with a more automatic implicit influence of spacing.

References

Abrahams, J. P., & Camp, C. J. (1993). Maintenance and generalization of object naming training in anomia associated with degenerative dementia. *Clinical Gerontologist*, 12(3), 57–72.

Armitage, S. G. (1946). An analysis of certain psychological tests used in evaluation of brain injury. *Psychological Monographs*, 60(1, Whole No. 277), 1–48.

Balota, D. A., & Duchek, J. M. (1989). Spreading activation in episodic memory: Further evidence for age-independence. *Quarterly Journal of Experimental Psychology*, 41(A), 849–876.

Balota, D. A., Duchek, J. M., & Paullin, R. (1989). Age-related differences in the impact of spacing, lag, and retention interval. *Psychology and Aging*, 4, 3–9.

Balota, D. A., & Faust, M. (2001). Attention in dementia of the Alzheimer's type. In F. Boller & S. Cappa (Eds.), *Handbook of neuropsychology* (2nd ed., Vol. 6, pp. 51–80). New York: Elsevier Science.

Balota, D. A., & Ferraro, F. R. (1996). Lexical, sublexical, and implicit memory processes in healthy young, healthy older adults, and in individuals with dementia of the Alzheimer's type. *Neuropsychology*, 10, 82–95.

Benton, A. L. (1963). The revised Visual Retention Test: Clinical and experimental applications. San Antonio, TX: Psychological Corporation.

- Berg, L. (1988). Clinical Dementia Rating (CDR). Psychopharmacology Bulletin, 24, 637–639.
- Berg, L., McKeel, D. W., Miller, P. J., Storandt, M., Rubin, E. H., Morris, J. C., et al. (1998). Clinicopathologic studies in cognitively healthy aging and Alzheimer disease: Relation of histologic markers to dementia severity, age, sex, and apolipoprotein E genotype. Archives of Neurology, 55, 326–335.
- Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA: MIT Press.
- Camp, C. J., Bird, M. J., & Cherry, K. E. (2000). Retrieval strategies as a rehabilitation aid for cognitive loss in pathological aging. In R. D. Hill, L. Backman, & A. Neely Stigsdotter (Eds.), Cognitive rehabilitation in old age (pp. 224–248). Oxford, England: Oxford University Press.
- Camp, C. J., Foss, J. W., Stevens, A. B., & O'Hanlon, A. M. (1996). Improving prospective memory task performance in persons with Alzheimer's disease. In M. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), *Prospective memory: Theory and applications* (pp. 351–367). Mahwah, NJ: Erlbaum.
- Camp, C. J., & Stevens, A. B. (1990). Spaced-retrieval: A memory intervention for dementia of the Alzheimer's type. *Clinical Gerontologist*, 10(1), 58-61.
- Carpenter, S. K., & DeLosh, E. L. (2005). Application of the testing and spacing effects to name learning. Applied Cognitive Psychology, 19, 619-636.
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (in press). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*.
- Crowder, R. G. (1976). *Principles of learning and memory*. Oxford, England: Erlbaum.
- Cull, W. L. (2000). Untangling the benefits of multiple study opportunities and repeated testing for cued recall. *Applied Cognitive Psychology*, 14, 215–235.
- Cull, W. L., Shaughnessy, J. J., & Zechmeister, E. B. (1996). Expanding understanding of the expanding-pattern-of-retrieval mnemonic: Toward confidence in applicability. *Journal of Experimental Psychology: Applied*, 2, 365–378.
- Estes, W. K. (1955). Statistical theory of distributional phenomena in learning. *Psychological Review*, 62, 369–377.
- Faust, M. E., Balota, D. A., & Spieler, D. H. (2001). Building episodic connections: Changes in episodic priming with age and dementia. *Neuropsychology*, 15, 626-637.
- Ferraro, F. R., Balota, D. A., & Connor, L. T. (1993). Implicit memory and the formation of new associations in nondemented Parkinson's disease individuals and individuals with senile dementia of the Alzheimer type: A serial reaction time (SRT) investigation. *Brain and Cognition*, 21, 163–180.
- Gabrieli, J. D. E., Vaidya, C. J., Stone, M., Francis, W. S., Thompson-Schill, S. L., Fleischman, D. A., et al. (1999). Convergent behavioral and neuropsychological evidence for a distinction between identification and production forms of repetition priming. *Journal of Experimental Psychology: General*, 128, 479–498.
- Glenberg, A. M. (1977). Influences of retrieval processes on the spacing effect in free recall. *Journal of Experimental Psychology: Human Learn*ing & Memory, 3, 282–294.
- Goodglass, H. & Kaplan, E. (1983a). Boston Diagnostic Aphasia Examination (3rd ed.). Philadelphia: Lea & Febiger.

- Goodglass, H., & Kaplan, E. (1983b). Boston Naming Test. Philadelphia, PA: Lea & Febiger.
- Grober, E., & Sliwinski, M. (1991). Development and validation of a model for estimating premorbid verbal intelligence in the elderly. *Journal of Clinical and Experimental Neuropsychology*, 13, 933–949.
- Hochhalter, A. K., Overmier, J. B., Gasper, S. M., Bakke, B. L., & Holub, R. J. (2005). A comparison of spaced retrieval to other schedules of practice for people with dementia. *Experimental Aging Research*, 31, 101–118.
- Hughes, C. P., Berg, L., Danziger, W., Coben, L. A., & Martin, R. L. (1982). A new clinical scale for the staging of dementia. *British Journal of Psychiatry*, 140, 566–572.
- Jacoby, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning and Verbal Behavior*, 17, 649–667.
- Karpicke, J. D., & Roediger, H. L. III (2005, May). Does expanding retrieval work? Poster session presented at the Annual Meeting of the Midwestern Psychological Association, Chicago.
- Landauer, T. K., & Bjork, R. A. (1978). Optimum rehearsal patterns and name learning. In M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 625–632). London: Academic Press.
- Light, L. L., & LaVoie, D. (1993). Direct and indirect measures of memory in old age. In P. Graf & M. E. J. Masson (Eds.), *New directions in cognition, development, and neuropsychology* (pp. 207–230.). Hillsdale, NJ: Englandiates.
- Logan, J. M. (2004). Spaced and expanded retrieval effects in younger and older adults. Unpublished doctoral dissertation, Washington University, St. Louis, MO.
- McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA work group under the auspices of the Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*, 34, 939–944.
- Morris, J. C. (1993). The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology*, 43, 2412–2414.
- Perry, R. J., & Hodges, J. R. (1999). Attention and executive deficits in Alzheimer's disease: A critical review. *Brain*, 122, 383–404.
- Rea, C. P., & Modigliani, V. (1985). The effect of expanded versus massed practice on the retention of multiplication facts and spelling lists. *Human Learning: Journal of Practical Research & Applications*, 4, 11–18.
- Spieler, D. H., & Balota, D. A. (1996). Characteristics of associative learning in younger and older adults: Evidence from an episodic priming paradigm. *Psychology and Aging*, 11, 607–620.
- Thurstone, L.L., & Thurstone, L.G. (1949). Examiner manual for the SRA Primary Mental Abilities Test. Chicago: Science Research Associates.
- Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80, 359–380.
- Verkoeijen, P. P. J. L., Rikers, R. M. J. P., & Schmidt, H. G. (2004). Detrimental influence of contextual change on spacing effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 796–800.
- Wechsler, D. (1955). Wechsler Adult Intelligence Scale (manual). San Antonio, TX: Psychological Corporation.
- Wechsler, D., & Stone, C.P. (1973). Wechsler Memory Scale (manual). San Antonio, TX: Psychological Corporation.

Received February 1, 2005
Revision received July 5, 2005
Accepted July 14, 2005