Journal of Clinical and Experimental Neuropsychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ncen19

Remediation of memory disorders: Experimental evaluation of the Spaced-Retrieval technique

Daniel L. Schacter [a], Susan A. Rich [a] & Michèle S. Stampp [a]

[a] University of Toronto ,

Published online: 04 Jan 2008.

Journal of Clinical and Experimental Neuropsychology 0168-8634/85/0701-0079 $3.00
1985, Vol. 7, No. 1, pp. 79-96

Remediation of Memory Disorders: Experimental Evaluation of the Spaced-Retrieval Technique*

Daniel L. Schacter, Susan A. Rich, and Michèle S. Stampp
University of Toronto

ABSTRACT

Research concerning remediation of memory disorders has frequently been concerned with mnemonic techniques that demand a great deal of elaborative and effortful processing. The present study examines a relatively simple technique, known as *spaced retrieval*, in which patients are taught to retrieve information at increasingly long temporal intervals after initial presentation. Results indicated that the spaced-retrieval technique aided patients' learning of new information. There was also evidence of *learning to learn:* Two of the four patients who were studied learned to use the technique in the absence of explicit cues from the experimenter. Issues pertaining to the possible usefulness of spaced retrieval in everyday life are discussed.

It is well known that memory deficits are among the most common sequelae of various types of neurological dysfunction (Cermak, 1982; Schacter & Crovitz, 1977; Whitty & Zangwill, 1977). Not until recently, however, has there been empirical research that has examined the possibility that memory functions of brain-damaged patients can be improved by appropriate cognitive interventions. The general approach that has been taken in existing studies is to attempt to modify patients' encoding of to-be-remembered information by introducing a mnemonic strategy that is known to improve the performance of normal subjects. The most frequently used intervention is the centuries-old technique of *visual imagery mnemonics* (Yates, 1966): Patients are instructed to construct vivid visual

* This research was supported by a Special Research Program Grant from the Connaught Fund, University of Toronto, and by a post-doctoral fellowship from the Ontario Ministry of Health awarded to Susan Rich. We thank Elana Joram and Carol MacDonald for experimental assistance, and thank Endel Tulving for useful comments concerning an earlier draft of the manuscript. Susan Rich is now at the West End Crèche, Toronto, and Michèle Stampp is now at Sunnybrook Hospital, Toronto.
Requests for reprints should be addressed to Daniel L. Schacter, Department of Psychology, University of Toronto, Toronto, Canada M5S 1A1.

Accepted for publication: April 25, 1984.

images that represent to-be-remembered information in a striking and distinctive format. Several studies have found that the performance of memory-impaired patients, like that of normals, can be improved by imagery mnemonics (Cermak, 1975; Crovitz, 1979; Crovitz, Harvey, & Horn, 1979; Gasparrini & Satz, 1979; Jones, 1974; Kovner, Mattis, & Goldmeier, 1983; Lewinsohn, Danaher, & Kikel, 1977; Patten, 1972; Wilson, 1981, 1982), although not all patients benefit from them (Crovitz et al., 1979; Jones, 1974). In addition to imagery mnemonics, strategies such as organization and chunking (Gianutsos & Gianutsos, 1979) and verbal labelling (Cermak, Reale, & DeLuca, 1977) have also been used with some success.

The foregoing studies may constitute a useful beginning for memory remediation research, but the strategies that they have employed are all characterized by a significant problem. Techniques such as imagery and organization require a great deal of cognitive effort to be used effectively by patients in their everyday lives; they require a kind of elaborative processing that most brain-damaged amnesics are unable to achieve spontaneously (Baddeley, 1982; Butters & Cermak, 1980). Indeed, one researcher has pointed out that the cognitive demands of imagery mnemonics are so extensive that he no longer uses them in his own daily activities (Cermak, 1980). However, a major shortcoming of existing research is that none of the studies that have employed imagery and organizational mnemonics have provided convincing evidence that patients use these techniques spontaneously either in the laboratory or in their everyday lives. Because of the extensive cognitive resources that are required to use imagery and organizational mnemonics spontaneously, it may prove difficult to teach patients to use these techniques on their own.

Are there any techniques that can aid mnemonic function without making excessive demands on patients' cognitive resources? One possibility has been suggested recently by Schacter (1980). Noting the substantial cognitive effort required to use imagery mnemonics, he suggested that a much less demanding technique known as *spaced retrieval* or *retrieval practice* (Bjork, 1979; Landauer & Bjork, 1978) might prove helpful for patients with organic memory disorders. Spaced retrieval is based upon an important property of the memory system: The act of retrieval exerts a powerful effect upon the subsequent memorability of a retrieved item. This idea was appreciated as long ago as the first decade of the 20th century by Richard Semon, an early memory theorist (see Schacter, 1982, pp. 187-189), but has only recently been explored by cognitive psychologists. Several studies have demonstrated that retrieval of an item on an initial test facilitates its later recall relative to initially untested items (Darley & Murdock, 1971; Modigliani, 1976; Whitten, 1978), and other experiments have revealed that retrieval of an item can be an even more effective aid to subsequent retention than an additional presentation of it (e.g., Izawa, 1969).

The technique of spaced retrieval represents an attempt to maximize the mnemonic benefits of recalling a previously studied item. It was first described in a study of normal college students reported by Landauer and Bjork (1978). They used a continuous paired-associate procedure in which study and test trials are intermixed. On the study trials, subjects saw a common name that was composed of a given name and a surname (e.g., Donald Williams). On the test trials, subjects were asked to recall the surnames in the presence of the given names. For each name, three successive tests followed an initial study trial; the test trials for a particular name were separated by study trials concerning other names. The critical feature of the task is that the temporal sequencing of the three tests was systematically varied. In the *uniform* condition, each of the three tests was separated by an equal number of intervening study trials. In the *expanding* condition, the three successive tests were separated by an increasingly large number of intervening study trials. In one experiment, for example, the first test immediately followed study, the second test occurred after three intervening study trials, and the third test was given after 10 intervening study trials. Landauer and Bjork found that any sequence of tests improved final recall performance with respect to a presentation-only condition. However, they consistently observed that the maximum benefit occurred in the expanding condition. They also found that an expanding test sequence yielded higher final recall than did a comparable sequence in which study trials concerning a particular name were administered in the same expanding sequence. Because the data suggest that the expanding pattern of retrievals is in some sense an optimal one, we use the term *spaced-retrieval technique* to refer exclusively to this temporal sequence of tests.

In addition to its potent effect on recall performance, one of the important features of spaced retrieval is that it seems to require little cognitive effort (Bjork, 1979). Acquisition of the technique consists of learning to attempt to retrieve to-be-remembered information at various intervals after presentation; there is no need to engage in the inventive and elaborative processing required by imagery and organizational strategies. For these reasons, we thought that it would be worthwhile to evaluate the possibility that spaced retrieval improves the mnemonic function of patients with memory disorders.

Before describing the experimental paradigm that was used to assess the effects of spaced retrieval, it is important to distinguish between two different aspects of patients' performance, both of which need to be considered in memory remediation research. On the one hand, it is necessary to determine whether a particular technique improves patients' ability to acquire or learn new information. Improvements in learning of new information are typically assessed by comparing indices of performance, such as proportion of items recalled or recognized, before and after introduction of a mnemonic technique. All of the previously cited studies of memory remediation have examined the extent to which a particular technique aids patients' ability to learn new information. However, a second aspect of performance that needs to be considered is whether patients can *learn to learn* – that is, whether they can acquire and retain knowledge of when and how to use a memory aid. The distinction between *learning* and *learning to learn* has long been known to experimental psychologists, but has received little attention so far in the literature concerning mnemonic training of brain-damaged patients.

The relevance of the foregoing distinction to the evaluation of the spaced-retrieval technique is highlighted by the results of a pilot study that we conducted in which two patients with relatively mild memory problems (one had incurred a left-sided CVA, and one did not have a firm neurological diagnosis) were exposed to the spaced-retrieval technique. Because patients frequently complain about their inability to remember the names and other characteristics of new acquaintances, the study focussed on patients' memory for characteristics associated with unfamiliar faces. Both patients were shown a series of faces and were told to try to remember a particular characteristic for each face, such as a name, occupation, or hobby. After a 90-s filled delay, they tried to recall the characteristic upon re-presentation of the face. Patients' recall performance was first assessed during six uninstructed baseline sessions. They were then introduced to the spaced-retrieval technique during eight training sessions. After exposure to the face-characteristic pair on each trial, patients were cued by the experimenter to retrieve the designated characteristic at three points during the 90-s delay that were separated by increasingly long temporal intervals (3 s, 10 s, and 30 s). After completion of these training sessions, a further six *assessment* sessions were completed, using similar materials and procedure, except that patients were not explicitly cued to engage in spaced retrieval.

The results indicated that both patients' recall of characteristics at 90-s delay improved substantially during the training sessions relative to performance during baseline sessions. Thus, as Landauer and Bjork (1978) observed with normal subjects, we found that spaced retrieval aided patients' ability to learn new information. However, patients' performance returned to baseline levels in the absence of explicit instructions to use spaced retrieval during the final six assessment sessions. This outcome indicates that patients had not yet *learned to learn* on their own.

In the research that is reported here, we explore in detail the extent to which spaced retrieval aids memory-disordered patients' ability to learn new information, and also examine whether patients can learn to use the technique on their own. These two issues were approached within the context of an experimental design that includes four principal phases: (1) a *baseline* phase in which patients' recall performance is measured prior to exposure to the spaced-retrieval technique; (2) a *cued-training* phase in which patients are exposed to spaced retrieval and are cued by the experimenter to initiate spaced retrieval; (3) a *self-cued training* phase in which the experimenter gradually withdraws the environmental cues to engage in spaced retrieval; and (4) an *assessment* phase that is in most respects identical to the baseline phase.

Two principal questions were of interest concerning the effect of spaced retrieval on patients' ability to learn new information. First, we wanted to know whether spaced retrieval can boost patients' memory for more than just a single item tested at a brief delay, as was observed in the pilot study. To investigate this issue, we tested patients on lists of items that varied in length, and compared proportion of items recalled during the baseline phase and the assessment phase. Since we studied patients whose level of memory performance varied widely, we had to calibrate appropriate list lengths individually for each patient in order to avoid ceiling and floor effects. Second, we wanted to ascertain whether spaced retrieval aids patients' ability to acquire information that is presented repeatedly across sessions. In the baseline, cued training, and assessment phases, different sets of experimental materials were used in each session. To evaluate learning of repeated materials, we exposed patients to the same set of items during all sessions of the self-cued training phase.

Quantitative evaluation of learning-to-learn effects was accomplished by recording the number of times patients required verbal prompts from the experimenter to initiate spaced retrieval during the assessment phase. The fewer prompts required by the patient, the greater the magnitude of the learning-to-learn effect.

METHOD

Subjects

Four patients participated in this study. Two of them contracted memory disorders secondary to a ruptured anterior communicating artery aneurysm (C.F. and W.H.), one became amnesic after a bout of viral encephalitis (H.G.), and in one case, there was no firm neurological diagnosis (L.H.). Patients L.H. and C.F. are characterized by mild memory problems, whereas patients H.G. and W.H. have severe memory problems. Information concerning patients' age, neurological diagnosis, and performance on the *Wechsler Adult Intelligence Scale - Revised* (WAIS-R) and *Wechsler Memory Scale* (WMS) is presented in Table 1.

Materials

One hundred and ninety-two black and white pictures of people's faces were presented on slides. The faces were culled from magazines and newspapers. They were projected by a Kodak Carousal slide projector that was connected to a stereo tape recorder. The tape recorder was wired so that it could drive the slide projector forwards and backwards at precisely timed intervals produced by prerecorded signals on the appropriate channel of the audio tape. For each of the 192 faces, a fictional given name, hometown, occupation, and hobby were generated by the experimenters. Examples of the fictional characteristics are presented in Table 2.

Design

The experiment can be conceptualized as an aggregate of four case studies, each of which is composed of a sequence of four consecutive phases: baseline, cued training, self-cued training, and assessment. Each phase was comprised of eight separate sessions during which patients studied faces and associated characteristics in lists that ranged in length from one to eight items. The lists were composed in the following manner. When a list consisted

Table 1

Characteristics of Memory-Disordered Patients

Patient	Diagnosis	Age	WAIS-R[a]	WMS[b]
C.F.	Aneurysm	39	102	90
L.H.	Uncertain	68	119	109
W.H.	Aneurysm	58	89	79
H.G.	Encephalitis	29	82	61

[a] Wechsler Adult Intelligence Scale - Revised (IQ)
[b] Wechsler Memory Scale (M.Q.)

Table 2

Examples of Names, Origins, Occupations, and Hobbies Used in the Experiment

Name	Origin	Occupation	Hobby
Mary	Charlottetown	dancer	embroidery
Douglas	Quebec City	actor	jogging
Cynthia	Vancouver	nurse	squash
Albert	Winnipeg	janitor	ham radio
Veronica	St. John's	stockbroker	art collector
Norman	Hamilton	steel worker	hunting
Jane	Montreal	newscaster	skating
Bernard	Toronto	engineer	skydiving

of four or fewer characteristics, all of the characteristics in the list were associated with one face. However, when a list consisted of more than four characteristics (in practice, either six or eight), an equal number of items related to each of two faces was used. For example, a list length of six would consist of three items related to one face and three items related to another face. For those list lengths that required less than the total of four items that were related to each face (i.e., list length of 1, 2, 3, or 6) the particular items that were used to make up a given list were chosen quasi-randomly. For example, a list length of two might be composed of a name and occupation on one occasion, hobby and origin on another, and an origin and name on still another occasion. The only constraint imposed upon construction of lists that did not use all four characteristics was that characteristics of each type (name, origin, occupation, hobby) appeared approximately equally often in each phase of the experiment; the design of the experiment did not permit complete counterbalancing across experimental phases. Assignment of individual faces and characteristics to the four experimental phases was done randomly.

Memory for characteristics associated with presented faces was assessed on three kinds of recall tests that were included in each of the four experimental phases: (1) *Item recall* refers to the test for each item that was administered at a 95-s delay; (2) *List recall* refers to the test that was conducted after all the characteristics in a particular list had been presented; and (3) *Session recall* refers to the test for all items of a given session; it was conducted after completion of each of the item and list recall tests of a particular session. In each of the foregoing tests, recall of characteristics was initially assessed by presenting a face and accompanying statement concerning the kind of information required (i.e., name, origin, hobby, or occupation). We refer to this type of test as *face-cued recall*. If patients failed to retrieve the appropriate characteristic on either the list or session tests, they were then provided with the initial two letters of the item. We refer to this test as *face and letter-cued recall*.

Procedure

Two sessions, separated by a ½ hour break, were held each time a patient visited the laboratory. All patients visited the laboratory twice a week for 8 weeks, and were tested individually.

Baseline. In all baseline sessions, patients were told that a face would be shown to them and that the experimenter would describe a characteristic of the person. The patients were instructed that they should try to remember the characteristic, and that they would be asked to recall it when the face slide was shown again about a minute and a half later. After presentation of a face and the first associated characteristic, patients completed cognitive tasks such as letter-search puzzles and various self-report inventories for the 95-s interval prior to the item test.

After 95 s had elapsed, the item test was administered by re-presenting the face slide and asking for recall of the appropriate characteristic. Ten seconds were allowed for recall. After completion of all item tests within a list in this manner, there was a further 30-s filled interval that was followed by the list test. Recall of characteristics was tested by exposing the previously seen faces in reverse order of presentation. The test of each characteristic associated with a particular face was completed before a new face was presented. If the characteristic was not retrieved, a cue that consisted of the first two letters of the characteristic was provided. Each successive list of items within a particular session was tested in the foregoing manner. After the last of the list tests, the session test for all characteristics was given. Order of testing was done in reverse order of initial presentation. If patients failed to retrieve the item on the face-cued recall test, an initial two-letter cue was provided. Within each baseline session, lists of items were presented and tested in an order determined by list length: Shorter lists were presented and tested before longer lists.

The first two baseline sessions were used to calibrate appropriate list lengths for each patient. The major purpose of the calibration procedure was to determine list lengths at which performance was above the floor and below the ceiling. In the first calibration session, all four patients were presented with and tested on lists of 2, 3, 4, and 6 items. In the second calibration session, patients were tested on either longer lists (8 items) or shorter lists (1 item), depending upon their performance in the first session. The shortest list length selected for each patient's subsequent baseline sessions was one item longer than the list length at which all items were consistently recalled during calibration sessions. For example, if a patient consistently recalled all of the items in a three-item list, the shortest list length that he or she would be assigned for the six subsequent baseline sessions would be four items. If a patient could not recall one-item lists consistently, then he or she was assigned a one-item list as the shortest list for baseline sessions. All patients were also assigned either two or three additional lists that contained progressively more items; the exact number of lists and items-within-lists depended upon patients' level of recall during calibration sessions. The four patients' performance differed during the calibration sessions, so the list lengths selected for the baseline and assessment phases varied among patients. Neither W.H. nor H.G., the patients characterized by severe memory disorders, could recall even one item consistently on the list test. Thus, for both of these patients, the shortest list length used during baseline and assessment was composed of a single item. Lists of two, three, and four items were also used during the baseline and assessment sessions of W.H. and H.G. Patient L.H. could recall one item consistently on the list test, but no more than that. Accordingly, her shortest list during baseline was composed of two items; longer lists, composed of three, four, and six items, were also used. Patient C.F. recalled three items consistently on the list test, so his shortest list was composed of four items. Because of restrictions in the available numbers of items, C.F. was assigned only two longer lists; they were composed of six and eight items.

Cued training. In this phase, patients were introduced to the spaced-retrieval technique. They were told that the face slide would reappear several times during the retention interval, and that they should pause from the interfering task when the slide appeared and try to state out loud the appropriate characteristic. After the initial presentation of the face and the statement of the to-be-remembered characteristic by the experimenter, the slide was removed for periods of 5, 10, and 20 s that were punctuated by 5-s intervals during which the patient tried to retrieve the to-be-remembered item. Following the third attempted retrieval of the item, there was a 40-s interval that was followed by the item test. The total time from the end of the initial presentation to the item test was 95 s. The item test and the list test were conducted in the same manner as in the baseline sessions.

Unlike the baseline phase in which lists of expanding length were used in each session, during cued training the same list length was used within a particular session. The length of the list used in the initial cued training session was equivalent to the shortest list length at which a particular patient did not consistently retrieve all characteristics on the list test during baseline. However, if a patient improved during treatment to a level at which he or she recalled all the items on the list test, a longer list was introduced in a subsequent session.

Self-cued training. During cued training, patients were signalled to initiate spaced retrieval both by the appearance of the face slide and by the verbal prompting of the experimenter. Both of these cues were withdrawn during self-cued training if and when patients exhibited two critical response patterns. If a patient spontaneously paused from the interfering task and looked at the face slide within 10 s of the appropriate time on 90% or more of the trials of a given session, then in the next session the face slide did not appear between initial presentation and the item test at 95-s delay. The patients did, however, continue to receive a verbal prompt to engage in spaced retrieval if they did not do so spontaneously within 10 s of the specified time at a particular interval. When patients attempted to retrieve the item spontaneously, this verbal prompt, too, was withdrawn, until and unless a patient failed to retrieve spontaneously on a subsequent trial; at this point, the verbal cue would be given again, for as many trials as it was needed.

The self-cued training phase also differed from the cued training phase in that the same lists of items were repeated throughout all eight self-cued training sessions. The target items were those that had been used in the final session of the cued training phase for each patient.

To evaluate between-session retention of the repeated items, each session test was administered again, in an identical manner, at the beginning of the next session. As in other experimental phases, two sessions were conducted each time the patient visited the laboratory. The first three visits were each separated by a 2-to-4 day interval. To provide information concerning possible long-term retention of repeated information, the final two sessions of the self-cued training phase were conducted after an interval of 12 days.

Assessment. The assessment phase was in most respects identical to the baseline phase. The same list lengths that had been used during baseline were used during assessment. None of the faces or characteristics included in the assessment phase had previously appeared at any point in the experiment, and different items were used in each session. Patients were told at the initiation of the assessment phase that they should continue to use the spaced-retrieval technique, but no face cues were provided at the spaced intervals. If patients failed to initiate retrieval at a particular interval, they were given a verbal prompt to do so. This procedure was adopted to permit estimation of the frequency with which patients initiated spaced retrieval spontaneously.

RESULTS

Comparison of baseline and assessment phases. Table 3 presents the proportion of items recalled during baseline and assessment phases on the item, list, and session tests. These data are collapsed across the six sessions of the baseline phase and the eight sessions of the assessment phase.

Consider first the data from the face-cued conditions. On the item tests that were administered at the 95-s delay, performance of all four patients improved from the baseline phase to the assessment phase. In the patients with severe problems, proportion of recalled items increased from .45 (H.G.) and .38 (W.H.) during baseline to ceiling levels during assessment. Patient L.H. showed a similarly large improvement of item recall. Patient C.F. also performed at a ceiling level during assessment, but the increase relative to baseline was somewhat smaller than in the other cases because of his high initial level of performance (.87). Statistical evaluation of differences between baseline and assessment phases was accomplished with a nonparametric test for comparison of two proportions that has been described by Bennett and Franklin (1954, pp. 611-615). This test revealed highly significant changes between baseline and assessment ($p < .01$) in all four patients.

The results of the list tests yielded a similar pattern of results. The data in Table 3 are collapsed across the various list lengths for each patient, but the major trends are observed at all list lengths in every patient (see below). Face-cued recall on the list test improved substantially from baseline to assessment in all patients. Particularly striking are the changes observed in patients H.G. and W.H., who recalled more than twice as many characteristics during the assessment phase than during the baseline phase. The Bennett-Franklin test revealed that proportion of items recalled during assessment significantly exceeded proportion recalled during baseline in each of the four patients (for H.G., W.H., and L.H., $p < .01$, for C.F., $p < .05$).

Table 3

Proportion of Characteristics Recalled on Item, List, and Session Tests During Baseline and Assessment Sessions

Patient		Type of Test					
		Item		List		Session	
		FC[b]	F+LC	FC	F+LC	FC	F+LC
L.H.	Baseline (90)[a]	.57	—	.37	.62	.13	.52
	Assessment (120)	.98	—	.68	.79	.20	.41
C.F.	Baseline (108)	.87	—	.36	.64	.20	.50
	Assessment (144)	1.00	—	.50	.63	.23	.53
H.G.	Baseline (60)	.45	—	.32	.67	.08	.48
	Assessment (80)	1.00	—	.71	.81	.19	.48
W.H.	Baseline (60)	.38	—	.23	.60	.12	.47
	Assessment (80)	.94	—	.64	.83	.23	.58

[a] The raw numbers of observations contributing to each cell in a particular condition are in parentheses. The numbers differ between baseline and assessment because the data are collapsed across six baseline sessions and eight assessment sessions.
[b] FC = Face-cued; F+LC = Face-and-letter cued.

The results of the session test yielded no more than suggestive evidence of improvement after training (Table 3). Proportion of items recalled by all patients increased numerically from baseline to assessment, but the observed changes were a good deal smaller than those reported for the item and list tests, and in no case did they approach statistical significance.

Consider next the data from the face and letter cued-recall test, also presented in Table 3. During baseline, presentation of initial letter cues on both the list and session tests produced significant ($p < .01$) increments of recall performance for all patients. A somewhat different pattern of results is observed during the assessment phase. On the session test, there was still a substantial advantage of face- and letter-cued recall over face-cued recall: All patients recalled significantly more characteristics with initial letter cues than without them ($p < .01$), and the magnitude of the advantage during assessment was similar to that observed during baseline. On the list test, however, only two patients (C.F. and W.H.) showed a statistically significant benefit of two-letter cues during the assessment phase ($p < .05$), and even in these cases, the magnitude of the cueing benefit was much smaller than the advantage that was observed during baseline. Remember, however, that all patients' face-cued performance on the list test improved significantly from baseline to assessment. This result suggests that the relatively small initial letter-cuing effect on the list test during assessment is attributable to the fact that more items on the list test were accessible to face-cued recall during assessment than during baseline. By contrast, on the session test – in which performance changed little from baseline to assessment – the benefit of cues was of approximately equal magnitude during both phases. Thus, it seems reasonable to suggest that use of the spaced-retrieval technique helped patients to gain access to items in the presence of face cues alone that previously required initial letter cues as well.

Consider next individual subject's session-to-session data for face-cued recall on the item tests and on the longest list length at which they were tested on the list test (Table 4). Two features of the data merit commentary. First, every patient showed substantial improvements on their longest list lengths. For example, on four-item lists, proportion of items recalled by the two severely amnesic patients, H.G. and W.H., improved, respectively, from .25 and .08 during baseline to .47 and .50 during assessment. Proportion of items recalled by L.H. increased from .28 to .48 on her longest list (six items) and C.F. improved from .25 to .36 on his longest list (eight items). Similarly, large baseline-to-assessment increments were observed at each of the shorter list lengths.

The second important point to note about the data in Table 4 is that the baselines are free of any systematic tendency to rise across sessions on either the item test or the list test. Nonrising baselines were also observed for every patient at each of the other list lengths that were used. Stable baselines indicate that observed improvements during the assessment phase can be attributed to the effects of spaced retrieval, and not to either generalized practice effects or to placebo effects produced by contact with or encouragement from the experimenter. The fact that patients' performance on the session test did *not* change from the baseline phase to the assessment phase provides further evidence against any interpretation of our positive results in terms of generalized practice effects or placebo effects.

Cued training. Patients began the cued training using list lengths that were one longer than the length at which they retrieved all items consistently on the list test during baseline. Patients L.H., H.G., and W.H. all began training with two-item lists; C.F., who exhibited higher list recall during baseline, began with a list length of four.

The effects of spaced retrieval on recall performance were observed early in the cued training. For example, on the item test at 95-s delay, all patients achieved perfect performance during the first cued training session. Item recall remained at ceiling for the duration of training, with the exception of four isolated cases of forgetting, distributed across patients and sessions. As is indicated by the data in Table 4, none of the patients had achieved perfect performance on the individual item test during any of the baseline sessions.

It is difficult to make straightforward comparisons between baseline and cued training performance on the list and session tests, because list lengths differed during the two experimental phases. Similarly, it is difficult to determine whether performance on the session test improved within the cued training phase, because

Table 4

Proportion of Characteristics Recalled on the Item Test and on the Longest List Test in Each Baseline and Assessment Session

Patient	Test type	Baseline 1	2	3	4	5	6	Assessment 1	2	3	4	5	6	7	8
L.H.	Item[a]	.80	.47	.50	.50	.56	.47	1.0	1.0	.90	1.0	1.0	1.0	1.0	1.0
	List	.50	.00	.33	.33	.50	.00	.50	.50	.33	.50	.33	.50	.50	.75
C.F.	Item[b]	.83	.89	.94	.75	.89	.83	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	List	.38	.13	.25	.25	.13	.38	.13	.50	.50	.50	.13	.25	.38	.50
H.G.	Item[c]	.40	.60	.40	.70	.30	.30	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	List	.25	.50	.00	.75	.00	.00	.25	.25	.25	.75	.50	.50	.50	.75
W.H.	Item[c]	.60	.30	.60	.40	.20	.10	1.0	.90	1.0	.70	1.0	1.0	.90	1.0
	List	.25	.00	.00	.25	.00	.00	.50	.50	.50	.25	.75	.25	.25	.50

[a] List length = 6
[b] List length = 8
[c] List length = 4

the total number of items on which patients were tested fluctuated across sessions. However, one point concerning the list test results during the cued training should be noted. Performance of three patients showed enough improvement (that is, it reached ceiling levels at a particular list length) for list length to be increased during training. L.H. and H.G., who began with a list length of two, moved up to a list length of three after their first and second respective training sessions. List length was increased to four following the fifth session for both patients. W.H., who also began with a list length of two, was exposed to a list length of three for all training sessions following the fourth one. Patient C.F., who began with a list length of four, was tested at this level throughout training because recall on the list test did not reach ceiling. Proportion of items recalled on the list test did, however, increase across sessions: During the first four cued training sessions, face-cued recall of C.F. was .45; it increased to .77 during the final four sessions.

Repetition of characteristics. The data in Table 5 portray each patient's performance on the face-cued recall tests for all characteristics that were administered at the beginning and end of each session during the self-cued training phase. Because of an experimental error, patient C.F. was given a new set of materials at the beginning of session six; his data are thus presented only for the first five cued-training sessions. There is a tendency for all patients to recall an increasingly higher proportion of the repeated characteristics across sessions and days. Consider, for example, the performance of the severely amnesic patient H.G. At the beginning of the first session, H.G. did not recall any of the characteristics that had been presented 2 days earlier during the final cued training session. She demonstrated negligible levels of between-session retention during the next four sessions of self-cued training, but she did demonstrate increasing amounts of within-session retention. H.G. then showed some evidence of between-session retention when tested at the beginning of session six (.17), and at the end of session six, H.G.'s final recall had increased to .58. H.G. demonstrated considerable retention after the 12-day retention interval, recalling .33 of the characteristics at the beginning of the seventh session, and concluded the final session at an even higher level (.83). Patients C.F., L.H., and W.H. also demonstrated relatively consistent gains across sessions and days. All patients (except C.F., for whom appropriate data are not available) exhibited higher levels of recall after the 12-day interval than they had on the initial session.

Table 5

Prorportion of Repeated Characteristics Recalled on the Initial and Final Session Tests During the Self-Cued Training Phase

	Session and Condition															
	1				2				3				4			
Day	1		2		3		4		5		6		7		8	
Patient	I[a]	F	I	F	I	F	I	F	I	F	I	F	I	F	I	F
L.H.[b]	.00	.38	.19	.31	.13	.38	.19	.38	.25	.44	.50	.50	.13	.50	.56	.69
C.F.[b]	.06	.50	.31	.50	.00	.69	.56	.88	.56	.69	—	—	—	—	—	—
H.G.[c]	.00	.17	.00	.25	.00	.25	.08	.33	.00	.42	.17	.58	.33	.67	.58	.83
W.H.[d]	.00	.33	.33	.56	.11	.55	.33	.33	.56	.44	.11	.22	.11	.56	.46	.56

[a] I = Initial test; F = Final test
[b] Total of 16 characteristics tested.
[c] Total of 12 characteristics tested.
[d] Total of 9 characteristics tested.

Further information concerning repetition effects is provided by analyzing the fate of individual items across sessions. One question of interest concerns whether patients were able to retrieve *any* items consistently across sessions. The protocols of patients L.H., H.G., and W.H. reveal that there was one item that was recalled in all of the sessions in which they recalled at least one characteristic. C.F.'s protocol for those sessions in which he received repeated charactersitics contains four items that were retrieved whenever a session contained some evidence of recall. Most items, however, were retrieved somewhat inconsistently from session to session. This point can be illustrated by considering the probability that, when an item has been recalled in a particular session, it will not be recalled again in the next session. This conditional probability of *item loss* was relatively high in patients L.H. (.43), W.H. (.62), and C.F. (.67). Only H.G. showed relatively little item loss (.14), but this finding must be treated cautiously because it is based upon only seven observations. One consequence of patients' somewhat inconsistent session-to-session recall is that all of them retrieved a large proportion of the to-be-remembered items at least once during self-cued training. C.F., for example, retrieved .88 of the characteristics at least once, and a majority of items were retrieved on one or more occasions by L.H. (.75), W.H. (.67), and H.G. (.58).

Learning-to-learn. Learning to learn effects can be evaluated by examining the proportion of trials during the assessment sessions in which patients required verbal prompts from the experimenter to initiate spaced retrieval (the face slide was not re-presented at the critical intervals during these sessions). Both C.F. and H.G. required no verbal prompts during any of the eight assessment sessions. By contrast, L.H. required verbal prompting to initiate spaced retrieval on .37 of the assessment trials; the proportion of prompted trials remained relatively constant across the assessment sessions. Patient W.H. required an even higher proportion of verbal prompts (.65) during the assessment sessions and in this case, too, the proportion of prompted trials showed only minor fluctuations across sessions.

The foregoing pattern of results – greater learning-to-learn effects exhibited by C.F. and H.G. than by L.H. and W.H. – is also evident in the data from the eight self-cued training sessions. Patients C.F. and H.G. met the criterion for withdrawal of the face slide (i.e., spontaneous looking at the slide on 90% of the trials) during the first session of self-cued training. H.G. required verbal prompting for only one further session, and C.F. required verbal prompting for two more sessions. L.H. did not meet the 90% criterion until the fifth session, and still required verbal prompts in 30% of the trials in the next three sessions. W.H. failed to meet the 90% criterion for withdrawal of the face slide at any point during self-cued training.

DISCUSSION

In the present research, we tried to evaluate the extent to which the spaced-retrieval technique aided patients' learning of new information, and also examined whether patients could be instructed to use the technique on their own – that is, whether they could learn to learn. Let us now consider the implications of the observed patterns of data with respect to each of these issues.

Comparison of proportion of characteristics recalled during the baseline and assessment phases revealed clearly that spaced retrieval improved the performance of all four patients on the item tests and list tests. By contrast, none of the patients' performance on the session test improved after exposure to the spaced-retrieval technique. These results may indicate that spaced retrieval is useful only when the amount of to-be-remembered information is relatively small. Alternatively, it is possible that spaced retrieval is only effective when information is tested at relatively brief temporal intervals: The session test occurred at a longer delay after initial presentation than did either the item test or list test. These and other possibilities could be explored in future research. The important point with respect to possible application of the spaced-retrieval technique is that we are now able to specify conditions under which the technique *does* enhance recall performance of memory-disordered patients.

The finding that spaced retrieval facilitated recall of a repeated set of characteristics across sessions and days provided further evidence that the technique aids patients' learning of new information. This finding has possible practical implications, in that it suggests that repeated use of the spaced-retrieval technique may facilitate retention of discrete bits of information that are important for a patient to remember. Although little attention has been paid to the possibility of using memory aids to attack individual problems in a patient's everyday life, this approach clearly merits investigation (Schacter & Glisky, in press). It would be particularly valuable at the present time to seek evidence that appropriate use of spaced retrieval can yield relatively permanent retention of an item or a number of items in memory-disordered patients. Patients did show retention of repeated characteristics over a 12-day interval in the present study, but their performance was far from perfect, and we do not know for how much longer they had access to the acquired information.

It is possible to argue, of course, that the positive results we did obtain with respect to learning of new information are not attributable specifically to the expanding pattern of retrievals that was used, but rather to the occurrence of repetition, independent of any particular pattern. We cannot address this issue purely on the basis of our data, because we did not include a condition in which there was nonspaced repetition of to-be-remembered information. Landauer and Bjork (1978) demonstrated that the expanding pattern of retrievals used in our experiment facilitates subsequent recall of normal subjects more than does non-spaced repetition, and we have no reason to believe that this is not the case in memory-disordered patients. However, the question of whether the mnemonic facilitation observed in our study is attributable to the specific pattern of spaced retrievals, or "merely" to repetition, is not a terribly important one. Our principal goal is to teach patients to benefit from the effects of repetition under conditions in which the environment cannot supply the needed repetition, such as when they try to remember the name of a new acquaintance. It is desirable, of course, to make use of the most effective pattern of self-administered repetitions, and available knowledge suggested to us that the spaced-retrieval technique meets this criterion. Evaluation of the effectiveness of other schedules of self-administered repetitions could be investigated in future studies of memory-disordered patients.

In addition to the effects on learning, we also found some evidence of learning to learn: Patients C.F. and H.G. used the spaced-retrieval technique in the absence of explicit cues throughout the assessment phase. W.H. and L.H., however, required periodic verbal prompts from the experimenter. In view of the fact that neither of the patients in our pilot study demonstrated any learning-to-learn effects, we think that it is reasonable to attribute the positive outcomes observed in the cases of C.F. and H.G. to the inclusion of the self-cued training phase. We also think that it is reasonable to hypothesize that with further self-cued training, W.H. and L.H., too, might demonstrate spontaneous use of spaced retrieval.

Demonstration of some effects on both learning and learning to learn brings us closer to the ultimate goal of any research concerning memory remediation: description of an effective technique that can be used by patients on an ongoing basis in their everyday lives. We have not, of course, provided evidence that spaced retrieval can help patients in their everyday lives, and neither do we know of any convincing evidence of long-term use of a mnemonic strategy in everyday life. What the present results do indicate is that spaced retrieval is sufficiently simple for some patients to use it spontaneously in the laboratory. It is entirely conceivable that patients will require explicit training in everyday contexts to recognize those situations in which it would be helpful to use the spaced-retrieval technique. Perhaps when we understand more about the problems faced by memory-disordered patients in natural environments (e.g., Schacter, 1983; Zola-Morgan & Oberg, 1980), we will be in a better position to design and implement techniques that can be used to combat the memory failures that plague patients in their everyday lives.

REFERENCES

Baddeley, A. D. (1982). Amnesia: A minimal model and an interpretation. In L. S. Cermak (Ed.), *Human memory and amnesia* (pp. 305-330). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Bennet, C. A.. & Franklin, N. L. (1964). *Statistical analyses in chemistry and the chemical industry.* New York: Wiley.

Butters, N., & Cermak, L. S. (1980). *Alcoholic Korsakoff's syndrome.* New York: Academic Press.

Bjork, R. A. (1979, October). Retrieval practice. Presented to *Conference on Developmental and Experimental Approaches to Memory,* Ann Arbor.

Cermak, L. S. (1975). Imagery as an aid to retrieval for Korsakoff patients. *Cortex, 11,* 163-169.

Cermak, L. S. (1980). Imagery and mnemonic training. In L. W. Poon, J. L. Fozard, L. S. Cermak, K. L. Arenberg, & L. W. Thompson (Eds.), *New directions in memory and aging* (pp. 507-510). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Cermak, L. S. (Ed.) (1982). *Human memory and amnesia.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Cermak, L. S., Reale, L., & DeLuca, D. (1979). Korsakoff patients' nonverbal and verbal memory: Effects of interference and mediation on rate of information loss. *Neuropsychologia, 15,* 303-310.

Crovitz, H. F. (1979). Memory retraining in brain-damaged patients: The airplane list. *Cortex, 15,* 131-134.

Crovitz, H. F., Harvey, M. T., & Horn, R. W. (1979). Problems in the acquisition of imagery mnemonics: Three brain-damaged cases. *Cortex, 15,* 225-234.

Darley, C. F., & Murdock, B. B. (1971). Effects of prior free-recall testing on final recall and recognition. *Journal of Experimental Psychology, 91,* 66-73.

Gasparrini, B., & Satz, P. (1979). A treatment for memory problems in left hemisphere CVA patients. *Journal of Clinical Neuropsychology, 1,* 137-150.

Gianutsos, R., & Gianutsos, J. (1979). Rehabilitating the verbal recall of brain-injured patients by mnemonic training: An experimental demonstration using single-case methodology. *Journal of Clinical Neuropsychology, 1,* 117-135.

Izawa, C. (1969). Comparison of reinforcement and test trials in paired-associate learning. *Journal of Experimental Psychology, 81,* 600-603.

Jones, M. K. (1974). Imagery as a mnemonic aid after left temporal lobectomy: Contrast between material-specific and generalized memory disorders. *Neuropsychologia, 12,* 21-30.

Kovner, R., Mattis, S., & Goldmeier, E. (1983). A technique for promoting robust free recall in chronic amnesia. *Journal of Clinical Neuropsychology, 5,* 65-71.

Landauer, T. K., & Bjork, R. A. (1978). Optimum rehearsal patterns and name learning. In K. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 625-632). New York: Academic Press.

Lewinsohn, P. M., Danaher, B. G., & Kikel, S. (1977). Visual imagery as a mnemonic aid for brain-damaged persons. *Journal of Consulting and Clinical Psychology, 45,* 717-723.

Modigliani, V. (1976). Effects on a later recall by delaying initial recall. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 609-622.

Patten, B. M. (1972). The ancient art of memory. *Archives of Neurology, 26,* 25-31.

Schacter, D. L. (1980, January). Imagery mnemonics, retrieval mnemonics, and the closed head injury patient. Presented to *International Neuropsychological Society,* San Francisco.

Schacter, D. L. (1982). *Stranger behind the engram: Theories of memory and the psychology of science.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Schacter, D. L. (1983). Amnesia observed: Remembering and forgetting in a natural environment. *Journal of Abnormal Psychology, 92,* 236-242.

Schacter, D. L., & Crovitz, H. F. (1977). Memory function after closed head injury: A review of the quantitative research. *Cortex, 13,* 150-176.

Schacter, D. L., & Glisky, E. (in press). Memory remediation. In B. Uzzell (Ed.), *Clinical neuropsychology of intervention.* Holland: Nijhoff.

Whitty, C. W. M., & Zangwill, O. L. (Eds.) (1977). *Amnesia.* London: Butterworths.

Whitten, W. B. (1978). Initial-retrieval "depth" and the relative recency effect. *Memory & Cognition, 6,* 590-598.

Wilson, B. (1981). Teaching a patient to remember people's names after removal of a left temporal tumour. *Behavioral Psychotherapy, 9,* 338-344.

Wilson, B. (1982). Success and failure in memory training following a cerebral vascular accident. *Cortex, 18, 581-594.*

Yates, F. (1966). The art of memory. Chicago: University of Chicago Press.

Zola-Morgan, S. M., & Oberg, R. G. E. (1980). Recall of life experiences in an alcoholic Korsakoff patient: A naturalistic approach. *Neuropsychologia, 18,* 549-557.