Journal Title: Memory & Cognition

Volume: 10
Issue:
Month/Year: 1982
Pages: 405-412

Article Author: Smith, S.M.

Article Title:). Enhancement of recall using multiple environmental contexts during learning.

Status: Faculty
Phone: (979) 845-2509
E-mail: stevesmith@neo.tamu.edu

Name: Steven Smith

Pickup at Evans

Address:
Psychology Department
MS# 4235
College Station, TX 77843

Memory & Cognition
1982, Vol. 10(5), 405-412

Enhancement of recall using multiple environmental contexts during learning

STEVEN M. SMITH
Texas A&M University, College Station, Texas 77843

Distributing the presentation of sublists of words into multiple learning rooms produced better free recall scores than a single learning room condition for subjects who were given a comprehensive recall test in a new environment. No such effects occurred on recognition or list differentiation tests in Experiment 2, implying a retrieval explanation rather than one relying upon learning or list differentiation effects. Experiment 3 found that the contextual dependence of recall (i.e., recall tested in a learning context is better than recall tested in a new context) was nullified by using multiple learning rooms, rather than a single room for input. The data are consistent with an explanation that states that the multiple learning rooms become associated with the different sublists during learning and subsequently act as memory landmarks that guide the course of retrieval.

Removal of contextual information has been shown to decrease recall performance relative to testing recall under conditions in which the learning context has been reinstated. This has been the case whether "context" has been defined as environmental surroundings (e.g., S. M. Smith, Glenberg, & Bjork, 1978) or as an internal pharmacological state (e.g., Eich, Weingartner, Stillman, & Gillin, 1975). Material studied by subjects in an experiment becomes connected in memory with information representing situational factors associated with the learning task, such that the reinstated context serves as a memory cue (S. M. Smith, 1979). Potentially confounding factors such as test room familiarity (S. M. Smith, 1979) or physical disruption (Godden & Baddeley, 1975) have been ruled out as alternate explanations of context-dependent findings.

The deleterious effects of testing memory in a context different from that in which learning occurred can be alleviated in at least two ways. One method is to have the changed-context subjects recall the setting in which learning took place and use the recalled environment as a retrieval cue (S. M. Smith, 1979). Another method is to test memory with a recognition test rather than free recall (Eich, 1980; S. M. Smith et al., 1978). These methods have at least two qualities in common; they are both procedures invoked after learning has already occurred, and both reduce the importance of having perceptually reinstated contextual cues at the time of testing. A recognition test provides excellent memory cues (copy cues of list words) for the subject, making general environmental cues relatively less useful. The context-recall procedure induces subjects to generate their own environmental retrieval cues, making experimenter-supplied environmental information relatively less important as a recall aid.

The present study discovered another method for eliminating contextual cue dependence, but through input manipulations rather than procedures employed at the time of testing. The basic phenomenon examined in the present study is the finding that subjects tested in a new environment will recall more if learning occurs in multiple environmental settings rather than a single environment. The lack of a consistent contextual referent for material memorized in multiple environmental contexts might make subjects less likely to rely upon the ambient test environment to enhance recall. The effects of contextual reinstatement (or contextual change), therefore, might not be seen if multiple learning rooms are used.

Environmental variations during learning can improve memory for a repeated list of words. S. M. Smith et al. (1978) found that when testing occurred in a new room, a list of words was recalled better if repetitions of the list during learning took place in two different rooms rather than in one room. Theoretically, this finding can be interpreted as showing that the organizational variability achieved through environmental changes allows flexibility in the subject's retrieval, so that a failure to retrieve an item stored in one mnemonic organization leaves the subject a second chance to retrieve the item while searching a second subjective organization. Organizational constancy does not provide this retrieval flexibility, as the subject has only one organization of list items to search.

This research was conducted at Texas A&M University. Thanks are due to Art Glenberg and Ernst Rothkopf for their helpful comments concerning the research, to J. B. Francks, Joy Kinney-Green, and Richard Gunn, who conducted the experiments, and to Robert A. Bjork, Eric Eich, and anonymous reviewers for comments on an earlier version of this manuscript. Requests for reprints should be sent to Steven M. Smith, Department of Psychology, Texas A&M University, College Station, Texas 77843.

0090-502X/82/050405-08$01.05/0

Closely related to this issue is the finding that memory improves as the number of subdivisions or categories in a list increases (e.g., Cohen, 1966; Mandler & Pearlstone, 1966). Although there is a tremendous difference between an experimental room and a semantically meaningful category, it is conceivable that both types of information could serve to subdivide a list of words and make it easier to recall.

Experiment 1 tested the hypothesis that when subjects are tested in a new environment, recall of multiple lists of words should be better for subjects who are presented the lists in multiple rooms, as compared with subjects who have all lists presented in the same environment. The effects of the length of the interval between lists, as well as the length of the retention interval, were also examined within this paradigm. Experiment 2 compared the effects of the input room manipulations on recall vs. recognition memory tasks. Finally, Experiment 3 examined the contextual reinstatement effect comparing single-room vs. multiple-room learning conditions.

EXPERIMENT 1

Subjects in Experiment 1 were given four lists of words, and they were given a single comprehensive free recall test afterward. Both environmental and temporal factors were manipulated. Of primary interest in this experiment was the effect of the number of input rooms: one, two, or four. The prediction was that memory tested in a new room would improve as the number of input rooms increased.

The temporal factors manipulated included the time between lists (interlist interval, or ILI) and the time between the fourth list and the free recall test (retention interval, or RI). Assuming that some type of temporally related memory cue is used to facilitate recall, it was hypothesized that such cues would be less available with a longer RI. Dependence upon environmental cues, then, should increase as the RI increases, leading to the prediction that the effects of contextual manipulations should be stronger with a longer RI. In regard to the ILI variable, it seemed possible that a long interval between lists could induce differentiation of lists in memory, thus having an effect similar to that caused by a room change. If so, then the effects of the room variations should be strongest at the shortest ILI, where temporal factors alone are not likely to induce list differentiation at the time of storage.

Method

Subjects. Ninety-six Texas A&M undergraduate volunteers assigned randomly to the various experimental groups were given course credit points for 1 h of participation in the present experiment. Experimental sessions were conducted in groups of two to four subjects at a time.

Procedure. Subjects were asked to memorize to the best of their ability four lists of aurally presented words. The lists of words were spoken at 3-sec intervals by a male voice on a cassette tape recorder. After each list, subjects were moved to a corridor and were asked not to discuss the experiment among themselves during each of these delays. After the interval following the fourth list of words, all subjects were taken to a new room and were given first a free recall test and then a recognition test. Subjects were given 8 min to write down all the words they could recall in any order, and they were encouraged to write down any words of which they were not certain. At the end of the free recall test, a recognition test containing 100 new words and 100 old words was given. Ten minutes were given for subjects to circle exactly 100 words that they felt were old list words. Subjects who recognized fewer than 100 were asked to guess until 100 were circled, and those recognizing more than 100 eliminated responses until only 100 remained.

Design. Four different between-subjects variables were manipulated in a 3 by 2 by 2 by 2 design, giving a total of 24 groups of subjects. Of primary interest was the number of input rooms (NR), which was one room (subjects returned to same room for each list), two rooms (first two lists given in one room, next two lists in a second room), or four rooms (subjects moved to a different room for each list). The ILI, or the time between lists, was either short (30 sec) or long (5 min). The RI or the time between the end of the fourth list and the beginning of the free recall test, was either short (30 sec) or long (5 min). Finally, two sets of words, Set A and Set B, were used. For one-half of the subjects, Set A comprised the four lists of memorized words and Set B served as distractors on the recognition test, and for the other half of the subjects, the sets were reversed.

Materials. The word sets were made up of four- and five-letter high-frequency English nouns selected from the Kučera and Francis (1967) word norms using the criterion that each word's frequency be greater than 50 per million. The 200 nouns were randomly divided into two sets of 100 (Sets A and B), and those sets were each subdivided into four lists containing 25 words apiece. The recognition test, which was identical for all subjects, consisted of a single page with all 200 words randomly ordered.

Rooms. Five rooms in the psychology department at Texas A&M University were used in the experiment. The rooms were selected to be as perceptually distinct as possible, each differing from the others in size, illumination, locations, decorations, and objects and apparatus placed in each room.

Results

Free recall. Two separate 3 by 2 by 2 by 2 (NR by ILI by RI by list) ANOVAs were calculated with the free recall data, one for the total number recalled per subject and another for the total amount of clustering by sublists for each subject.

The ANOVA using total number of words recalled found a significant effect for the NR variable $[F(2,72) = 3.96, MSe = 83.57, p < .02]$. The means for the one-room, two-room, and four-room groups, displayed in Table 1, showed the predicted effect, with greater recall scores associated with those groups given more input rooms. The two-room group recalled 1.94 more words

Table 1
Comparison of Mean Number of Words Recalled, Clustering Scores, and Recognition Hits in Experiment 1 as a Function of Number of Input Rooms

Measure	Number of Input Rooms		
	One	Two	Four
Number Recalled	19.9	21.9	26.2
Clustering	32.5	41.3	48.3
Recognition Hits	66.4	65.6	67.1

Note–Maximum possible was 100 for recall and recognition.

than the one-room group, an average improvement of 10% of the one-room group's total recall, whereas the four-room group recalled 6.28 more words than the one-room group, an average of 31% better recall.

The total number of words recalled was also significantly affected by the RI [$F(1,72) = 3.98$, $MSe = 83.57$, $p < .05$], with an average of 24.52 words recalled by the group with a 30-sec RI as compared with 20.83 words for the 5-min RI group. Although a longer RI produced worse recall, better recall was found for the groups given a longer ILI [$F(1,72) = 4.17$, $MSe = 83.57$, $p < .04$]. The 5-min ILI group recalled an average of 24.58 words, with 20.77 words recalled by the 30-sec group.

The interaction of NR by ILI by RI is shown graphically in Figure 1. Although the interaction did not even approach significance [$F(2,72) = .50$], it is clear from the figure that the groups given both a short ILI and a short RI did not contribute at all to the NR effect: All three groups had identical means.

The only other significant effects in the ANOVA using total words recalled were three interactions involving the list factor (a counterbalancing variable): ILI by List [$F(1,72) = 5.00$, $p < .05$], NR by RI by List [$F(2,72) = 4.10$, $p < .05$], and NR by ILI by RI by List [$F(2,72) = 5.58$, $p < .01$]. Systematic effects of the two counterbalancing lists were not apparent in an examination of the data, and the effects have no particular relevance to the theoretical issues considered in the present study.

An ANOVA using clustering scores was also calculated. Each subject's total clustering score was figured as the total number of adjacencies in the subject's

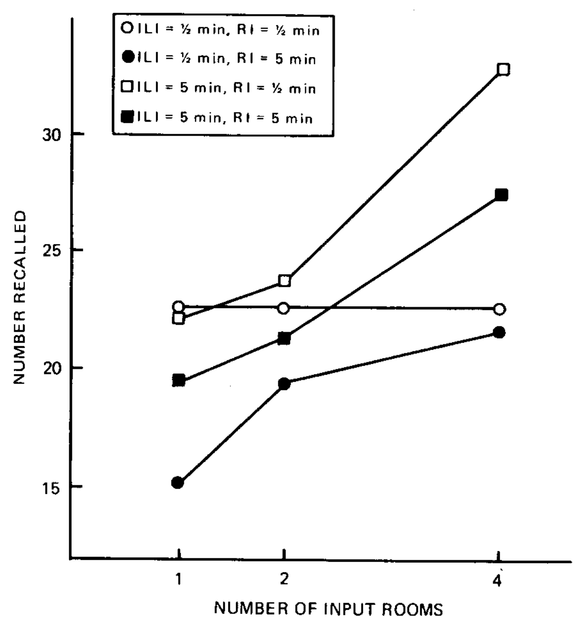

Figure 1. Mean number of words recalled as a function of number of input rooms, interlist interval (ILI), and retention interval (RI) in Experiment 1. Maximum possible was 100.

recalled words divided by the total number of possible adjacencies, given the number of words recalled from each sublist. Two words from the same sublist recalled consecutively were counted as an adjacency. The total number of possible adjacencies was calculated as the total number of words recalled minus the number of sublists from which at least one word was recalled.

The ANOVA using clustering scores found that NR had a very powerful effect [$F(2,72) = 9.0$, $MSe = 892.69$, $p < .001$], just as it had when total recall was used as a dependent measure. As shown in Table 1, clustering increased sharply with increasing NRs, with the clustering score of the four-room groups nearly 1.5 times that of the one-room group. The only other significant effect was the interaction of ILI by RI [$F(1,72) = 4.33$, $MSe = 892.69$, $p < .05$], which indicated that with the short ILI, a short RI produced more clustering than a long one, whereas RI had no effect when the ILI was long.

Recognition. An ANOVA was computed using number of hits per subject from the recognition memory test. Even though the error variance was relatively low in this ANOVA ($MSe = 88.60$), there were no significant effects found in the analysis. The mean number of hits for the one-room, two-room, and four-room groups are shown in Table 1. No hint of an effect can be seen in those means. When broken down according to the various ILI and RI conditions, the only groups showing the same NR trend seen in the free recall scores were the groups given a 5-min ILI and a 5-min RI. For those groups, the one-room group scored an average of 63.50 hits, compared with 70.75 hits for the four-room group. All other comparisons showed, if anything, a slight trend in the opposite direction, with fewer rooms producing better recognition.

Discussion

Although prior studies have shown facilitative effects of contextual cuing of memory, the present experiment has shown a somewhat different effect of environmental context, namely, its use at input to induce multiple organizations and subdivisions of a large set of information in memory. A categorized or subdivided list of otherwise unrelated items is recalled better than an uncategorized or undivided list (e.g., Puff, 1970), and the words tend to be clustered more by categories when they are recalled. The present results indicate that the environmental contexts in this study acted as organizational cues for the word lists and that those cues were used to guide the retrieval of the word lists. Even though recognition was not affected by the room manipulations, the two retrieval measures (free recall and clustering) were strongly improved by increasing NRs.

The effects of the temporal variables are not as easily evaluated as the effects of NR. The finding that a longer RI made recall worse is not surprising. That a longer ILI made recall better, however, is more interesting, since longer ILIs also have the effect of making the overall RI longer. Recall gains with longer ILIs could be caused by greater separation of sublists in memory, or, perhaps, by creating greater list differentiation. Separation of sublists in memory by increasing the ILI could functionally have the same effect as a room change, creating additional (temporal) retrieval cues for a longer ILI.

Empirically, it is interesting to note that the groups given both a 30-sec ILI and a 30-sec RI did not show any effect of NR. Follow-up experiments looking for this phenomenon should, therefore, use intervals longer than 30 sec.

EXPERIMENT 2

The free recall data in Experiment 1 are clear, in that they demonstrate the predicted improvement in recall with an increasing NR. The prediction of this phenomenon and the interpretation of it were based upon an explanation involving retrieval that is augmented by environmental cues stored in memory. Two alternate explanations, however, are also possible: a list differentiation hypothesis and a learning hypothesis.

The learning hypothesis states that more is learned in the multiple-room conditions than in the one-room condition, thus allowing for better recall performance. This explanation, for example, could be based upon a mechanism such as increased arousal or attention during learning caused by the change of environments.

The list differentiation hypothesis states that the room changes provide a mechanism for better separation in the storage of the different lists in memory. This could be achieved by the association of different contextual markers (Anderson & Bower, 1972) with the items in different lists, thus leading to better list differentiation, and concomitantly improving recognition memory (Anderson & Bower argued that recognition and list differentiation involve essentially the same processes). Since recognition is an important subprocess in generation-recognition models of recall, it seems possible that recall could be enhanced through a mechanism such as improved recognition caused by additional contextual list markers. To test this hypothesis more directly, a list differentiation test was administered in order to assess the subjects' ability to label items with their proper list membership.

Both alternate explanations, the learning hypothesis and the list differentiation hypothesis, make the prediction that increasing the NR should improve recognition memory performance. The learning hypothesis predicts that more learning allows more information to be recognized. The list differentiation hypothesis states that additional contextual list markers caused by room changes improve recognition memory. The retrieval hypothesis, however, does not predict an improvement in recognition with increasing NRs, stating that mutliple environments provide multiple retrieval cues for associated list words.

The recognition results from Experiment 1 fail to show the improvement predicted by the learning and list differentiation hypotheses, except for the groups given a 5-min ILI and a 5-min RI. It is possible that the improvement in recognition for this condition is a chance occurrence, or alternatively, it may be that recognition effects are dependent upon the longer intervals. Whatever explanation is true, however, it is clear that those recognition scores could have been contaminated by the previous free recall test. For this reason, recognition was tested first in Experiment 2, without any intervening memory tests. A long ILI and RI were used in Experiment 2 to maximize the likelihood of observing effects of input room manipulations on recognition memory.

The procedure of moving subjects into the corridor to wait between lists was a control method to assure that all subjects were equally disrupted between lists by physical movement. Strand (1970) believed that the physical disruption between lists could lead to better list differentiation, serving the same function and causing the same effects as room changes during learning. She used this notion to explain the reduction of interference found when original and interpolated lists are learned in different environments (Bilodeau & Schlosberg, 1951; Greenspoon & Ranyard, 1957). To test the possibility of disruption-induced list differentiation, a one-room undisrupted condition was added to the design of Experiment 2.

Method

Subjects. Forty-eight Texas A&M undergraduate volunteers assigned randomly to the various experimental groups were given course credit points for 1 h of participation in the present experiment. Experimental sessions were conducted in groups of two to four subjects at a time.

Design. NR, the variable of primary interest, was divided into three groups: a four-room group, a one-room group, and a one-room undisrupted group. The one-room and four-room groups were treated like the corresponding groups in Experiment 1. The one-room undisrupted group was treated like the one-room condition, except that the undisrupted subjects did not move into the hall to wait between lists; instead, they remained seated until all four lists had been presented. As in Experiment 1, two word sets (A and B) were used for different groups of subjects.

Procedure. The procedure prior to the memory test was identical to that used in Experiment 1. All intervals (ILI and RI) were 5 min. The first memory task was a two-alternative forced-choice recognition test with 100 pairs of test words. Each pair of words was randomly selected, one from Set A and one from Set B. Subjects were given 10 min to circle one member of each pair that they felt was an "old" list word.

The second test was a list differentiation test. Subjects were given a page containing a random ordering of the 100 originally studied words, and they were asked to indicate each word's list membership with a "1," "2," "3," or "4." Ten minutes were allowed for this test.

Materials. The list learning materials were identical to those used in Experiment 1. The recognition test consisted of 100 pairs of words; one word from Set A and one from Set B were randomly selected for each pair. Two different test pages were used for the list differentiation test, one for Set A and one for Set B. A test page contained a random ordering of all 100 words from a single word set.

Rooms. The rooms used in Experiment 2 were the same ones used in Experiment 1.

Table 2
Mean Number of Hits for Recognition and List Differentiation in Experiment 2 as a Function of Number of Input Rooms

Test Type	Number of Input Rooms		
	One*	One	Four
Recognition	81.6	79.5	79.1
List Differentiation	39.6	41.3	36.8

Note–*Maximum possible was 100.* **Undisrupted.*

Results and Discussion

The results of the recognition and list differentiation tests are shown in Table 2. Separate 3 by 2 (NR by list) ANOVAs were computed for the two measures, and both analyses confirm what is clear from Table 2; namely, that NR had no effect on either type of test [for the recognition test, $F(2,42) = .18$, $MSe = 104.3$; for the list differentiation scores, $F(2,42) = .30$, $MSe = 284.8$]. These results reaffirm what was found in Experiment 1, that the effect of NR is caused by a retrieval mechanism, rather than an improvement in item recognition or list differentiation. It should be noted that the list differentiation test may have been contaminated by the preceding recognition test; therefore, less importance might be attributed to the lack of list differentiation effects. The data also indicate that the room effect found in Experiment 1 was not caused by learning differences among those experimental groups, since recognition was equal for all groups.

EXPERIMENT 3

Subdividing a list of words and presenting them in multiple environments in Experiment 1 enhanced memory for subjects tested in a new room. Testing in a new environment has been shown to produce worse recall than testing in the context in which learning occurred, presumably because the learning context serves as a recall cue (S. Smith & Guthrie, 1924; S. M. Smith, 1979). The question asked in Experiment 3 is whether the input manipulation used in Experiment 1 will nullify the recall aid that is usually provided when the learning context is reinstated at the time of testing.

Some support for this notion was reported by Pan (1926), who presented incidental context words with paired associates. If a context word that was related to a response word was removed during recall rather than reinstated, recall was decreased. The decrement was diminished, however, if context words for pairs had been varied during learning. Although Pan's study used verbal context and repeated presentations of paired associates, as compared with environmental context and single presentations, his results are suggestive in regard to the present experiment.

The typical verbal memory experiment is conducted in a single laboratory setting, thus providing a unique set of environmental information that might later be used as a good retrieval cue for words memorized in that settng. If, on the other hand, material is presented in multiple learning environments, the subject can no longer rely upon a single set of information to serve as an environmental retrieval cue, and such a subject may be less likely to rely upon the ambient test environment to provide memory cues on a recall test. Subjects given multiple learning rooms, therefore, might not show the effects of contextual change/reinstatement. This possibility was tested in Experiment 3. Subjects were given either one or three learning rooms in which they studied three lists of words, and they were given a free recall test either in a totally new room (called the different-context, or DC, condition), or they were tested in a room in which learning occurred (same-context, or SC, condition). The design for these four major conditions (SC one room, DC one room, SC three rooms, DC three rooms) is shown in Table 3.

It is important to note that the SC three-room groups were tested in only one of their three learning rooms (the test room was counterbalanced among the three learning rooms). Those groups may have received less contextual cuing as compared with the SC one-room group. Therefore, it was necessary to also examine within-subjects comparisons for the SC three-room groups, comparing each subject's recall of the list learned in the test room with his recall of the other two lists of words. This comparison was similar to one used by S. M. Smith et al. (1978, Experiment 2), who found that paired associates learned in the test room were recalled better than those learned in another room.

Experiment 3 also provides a partial replication of the results found in Experiment 1, which showed that for subjects tested in a new room, recall was better for those given multiple input rooms than for those given only one learning room. Three lists and 5-min intervals were used for all subjects.

Method

Subjects. One hundred and twenty Texas A&M undergraduate volunteers assigned randomly to the different experimental groups were given course credit points for 1 h of participation in the present experiment. Each experimental session was conducted in groups of two to five subjects at a time.

Materials. Three lists were constructed, each consisting of 32 four- and five-letter common English nouns selected from the

Table 3
Rooms Used in Experiment 3 for Presentation of List 1, List 2, List 3, and Free Recall Test, Respectively

Test Context	Number of Input Rooms	
	One	Three
Same	AAA-A	ABC-A
	BBB-B	ABC-B
	CCC-C	ABC-C
Different	AAA-D	ABC-D
	BBB-D	ABC-D
	CCC-D	ABC-D

Kučera and Francis (1967) word frequency norms, using the criterion that words must be of a frequency of 50 per million or greater. The words were recorded on an audio cassette recorder by a male voice.

Procedure. The procedure was essentially identical to that used in Experiment 1, with a few exceptions. Instead of four lists, only three lists of 32 words were presented to subjects. All intervals (ILI and RI) were 5 min. Also, no recognition test was given, only a free recall test.

Design. For half of the subjects, the three lists were presented in three different rooms, Rooms A, B, and C. The other half of the subjects heard the lists all in the same room. One third of the one-room subjects heard the words in Room A, another third in Room B, and another third in Room C. As in Experiment 1, subjects in the one-room condition left the room and returned between lists.

Crossed with the NR variable was the test room variable with two levels: SC and DC. The DC subjects were tested in a new room, Room D. The SC subjects were tested in an "old" room, that is, a room in which learning occurred. The SC one-room groups were tested in their one and only learning room. One-third of the SC three-room groups were tested in Room A, one-third in Room B, and one-third in Room C (see Table 3).

Results and Discussion

The results of the four main conditions are shown in Figure 2. Clearly, the effect of the test context (SC vs. DC) was quite large in the one-room condition, whereas the same effect was nonexistent in the three-room condition. A 2 by 2 (test context by NR) ANOVA using subjects' free recall scores found a significant effect of test context [$F(1,116) = 3.99$, $MSe = 104.9$, $p < .05$], an effect caused primarily by the one-room groups. The interaction of test context with NR did not quite reach significance [$F(1,116) = 2.81$, $MSe = 104.9$, $p < .10$], apparently because of the large error variance. A Tukey's HSD test showed that the effect of the test context was significant for the one-room groups, but not for the three-room groups. For the one-room groups, the SC scores were an average of 42% higher than scores for the DC group. For the three-room groups, however, the difference between SC and DC recall was less than 3%. A robust effect of reinstatement vs. change of environmental context was shown in the typical one-room condition, but the effect was wiped out by varying the contexts at input.

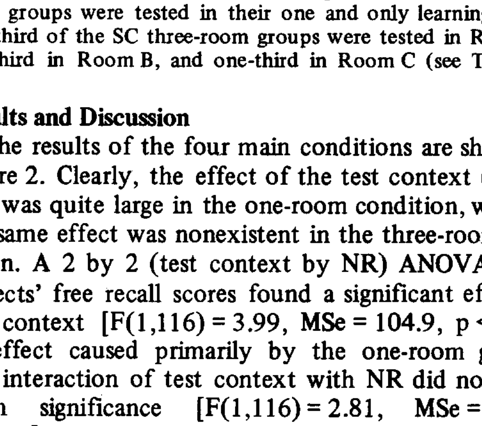

Figure 2. Mean number of words recalled in Experiment 3 as a function of test context and number of input rooms. Maximum possible was 96.

Table 4
Frequencies of SC Three-Room Subjects in Experiment 3 Beginning Recall With Items From List 1, List 2, and List 3

Counterbalancing Group	List Membership of First Item Recalled		
	List 1	List 2	List 3
ABC-A	9	0	1
ABC-B	5	4	1
ABC-C	5	3	2

Note—Maximum frequency for each row was 10 subjects.

A problem remains, however, in the use of the SC three-room group, in that only one of the three sublists may have received contextual cuing: namely, the list presented in the final test room. If such were the case, then at the very least, one would expect recall of the list presented in the test room (the SC list) to be recalled better than the lists presented in other rooms (DC lists) for the SC three-room groups. In fact, DC lists were recalled slightly better than the SC lists, with an average of 7.63 words on DC lists as compared with 6.63 words for SC lists. This result was unexpected, and it is somewhat at odds with Experiment 2 of S. M. Smith et al. (1978), but the reason for this discrepancy is unclear at this time. At any rate, it does not appear that recall in the SC three-room condition would have been improved even if it had been possible to cue those subjects with all three input environments.

An examination of output order for subjects in the SC three-room condition added little to the clarification of that group's performance (see Table 4). It seemed possible that subjects might begin recall with words from the list associated with their test room. There was a slight trend in that direction: More subjects in the ABC-A group began recall with a List 1 word than did subjects in the other two groups, more ABC-B subjects began with List 2 than did other groups, and more ABC-C subjects began with List 3 than did other groups. That slight trend, however, received no statistical support [for frequencies of subjects beginning recall with List 1, $\chi^2(2) = 1.69$; for List 2, $\chi^2(2) = 3.72$; and for List 3, $\chi^2(2) = .50$].

A comparison of the DC one-room group's recall with that of the DC three-room group is important to replicate the phenomenon noted in Experiment 1, in which it was found that for DC subjects, increasing the NR improved final free recall. Unfortunately, the present results are slightly ambiguous. The mean for the DC one-room group was 16.60 words, as compared with 20.63 words for the DC three-room group, an improvement of 24%. Although this effect is similar in magnitude and direction to that found in Experiment 1, the present effect was not significant, according to a

Tukey's HSD test. The potential reasons for this discrepancy are numerous, beginning with the fact that the error variance in Experiment 3 was relatively large. There are other differences between Experiments 1 and 3; the total number of words presented, number of lists, number of words per list, NR, and the time intervals used were a few of the factors that differed. Those factors responsible for the discrepant results are not immediately apparent. Even though the effect was not significant in Experiment 3, there was a clear trend in the results confirming the phenomenon seen in Experiment 1.

GENERAL DISCUSSION

Loss of contextual information, or its absence from a memory testing situation, is a major source of forgetting in a variety of circumstances (e.g., Abernethy, 1940; Eich, 1980; Godden & Baddeley, 1975; S. M. Smith, 1979). Operations designed to remedy this contextual dependence of memory have typically focused upon output manipulations; such operations include providing category cues (Eich et al., 1975), using recognition instead of a recall test (S. M. Smith et al., 1978), or asking subjects to generate their own contextual retrieval cues from memory (S. M. Smith, 1979). Instead of a "remedy" for contextual dependence in memory, however, the present study has examined a method for "immunizing" learned information against the negative effects of contextual changes, that is, performing operations at input rather than at output to make learned material less contextually bound.

The present experiments discovered that when recall was tested in a new environment, performance was better for subjects whose learning had been distributed among many contexts than for those who learned material all in one environment. In Experiment 1, it was shown that recall and clustering measures increased as NR increased. This same trend showed up again in Experiment 3, comparing one-room vs. three-room conditions, and although the effect did not reach significance, the size of the effect was similar to that observed in Experiment 1. For convenience, this phenomenon will be referred to as the multiple input context effect.

In Experiment 3, it was found that the multiple input context effect is more than a simple enhancement of memory; an increase in NR improved recall for subjects tested in a new context, whereas no improvement was seen for subjects tested in a reinstated environmental context. The multiple input context manipulation appears to have "immunized" learned material against the negative effects of contextual change.

A number of interpretations of the multiple input context effect are possible, although few explanations account for all of the evidence provided by the present set of experiments. In particular, most of the hypotheses cannot explain why the multiple input context effect is evidenced with a free recall test, but not with a recognition test, a result that was found in Experiments 1 and 2. For example, an explanation that states that more is learned in the multiple-room condition than in the single-room condition predicts similar results for recognition and recall tests. A hypothesis based upon enhanced list differentiation caused by room changes also predicts multiple input context effects on a recognition test and on a list differentiation test (neither of which was found). Likewise, recognition memory effects are predicted by a hypothesis that states that multiple room changes during learning serve to inure subjects to the negative effects of a room change at the time of testing. An adequate explanation of the phenomenon must handle the discrepancy between recall and recognition results.

The multiple input context effect resembles to some degree the improvement in recall noted when a set of words is subdivided into increasing numbers of categories or subjective organizations (e.g., Cohen, 1966; Mandler & Pearlstone, 1966). There are, however, some important differences between the results of the present experiments and the results of categorized or subjectively organized lists. First, the "some-or-none" principle of category recall (Cohen, 1966) states that if subjects recall any items at all from a category, then they are likely to recall a relatively fixed proportion of the entire category. Clearly, subjects in the present study did not simply recall a fixed proportion of a sublist, and only very rarely did subjects completely omit a sublist of words. Another difference between the two types of studies again concerns the recall vs. recognition memory differences with respect to susceptibility to the multiple input context effect. In the present study, only recall was affected and recognition was not. Using categorized lists, however, the number of categories used has been shown to affect not only recall, but recogniton, too (Mandler, Pearlstone, & Koopmans, 1969). Such discrepancies weaken the analogy of a categorized list as an explanation of the multiple input context effect.

A hypothesis that is consistent with all of the present results is one that states that the environmental context experienced during learning becomes associated with the learned material, and the context serves as a memory landmark. Such landmarks can be subsequently generated from memory and used to guide retrieval of the various subdivisions of list material. The more landmarks established during learning, the more retrieval starters will be available at the time of testing. Hence, the number of items recalled, as well as the amount of clustering by sublist should increase as more input rooms are used. Recognition memory, according to this explanation, need not be affected by the number of memory landmarks, since recognition processes are so strongly guided and controlled by the copy cues provided on the test. This landmark hypothesis differs from a category type of explanation because a landmark could serve as a retrieval starter even if it were associated with as few as one of the sublist items, whereas a category name would be associated with all of its exemplars.

Generating one's own contextual memory cues or landmarks as an initial step in recalling memorized material has been demonstrated to be a facilitative recall strategy (S. M. Smith, 1979). Operations performed during learning can provide the subject with memory cues that are useful even if those cues are not supplied by the experimenter at the time of testing. If, indeed, subjects in the multiple input room conditions in the present experiment used self-generated contextual retrieval cues to enhance recall, then it would appear that contextual cues were equally available whether subjects were tested in a new room or in one of the old learning rooms, since there was no difference in recall between the SC three-room and DC three-room conditions in Experiment 3.

The results of the manipulations of time intervals in Experiment 1 are by no means conclusive, but the fact that the multiple input context effect was totally absent in the condition in which brief ILIs and RIs were used indicates the sensitivity of this phenomenon to temporal conditions. Theoretically, the results might mean that when very recent events are remembered, less emphasis is given to recall strategies that employ contextual information, and when events are less recent, contextual cues are more likely to be employed. Empirically, it is important to note that a departure from the time intervals used in the present study could cause a failure of the present results to be replicated. At this point, however, it is not possible to specify the optimal conditions for obtaining the multiple input context effect.

Practial applications of the multiple input context effect seem possible, particularly in an educational setting. The fact that most classroom learning is intended to be used in situations outside the classroom exemplifies the problem of memory loss caused by changed contextual conditions. The multiple input room procedure appears in a sense to "immunize" learned material against the negative effects of a changed testing context, and it might be useful in a classroom setting. In fact, results of a study by S. M. Smith and Rothkopf (Note 1) indicate that the variation of input contexts benefited memory for information learned in a minicourse in statistics. More powerful applications of the effect may be possible when the phenomenon is better understood.

REFERENCE NOTE

1. Smith, S. M., & Rothkopf, E. Z. *Varying environmental context of lessons to compensate for massed teaching.* Paper presented at the convention of the American Educational Research Association (AERA) in New York, 1982.

REFERENCES

Abernethy, E. M. The effect of changed environmental conditions upon the results of college examinations. *Journal of Psychology*, 1940, **10**, 293-301.

Anderson, J. R., & Bower, G. H. Recognition and retrieval processes in free recall. *Psychological Review*, 1972, **79**, 97-123.

Bilodeau, I. M., & Schlosberg, H. Similarity in stimulating conditions as a variable in retroactive inhibition. *Journal of Experimental Psychology*, 1951, **41**, 199-204.

Cohen, B. A. Some-or-none characteristics of coding behavior. *Journal of Verbal Learning and Verbal Behavior*, 1966, **5**, 182-187.

Eich, J. E. The cue-dependent nature of state-dependent retrieval. *Memory & Cognition*, 1980, **8**, 157-173.

Eich, J. E., Weingartner, H., Stillman, R. C., & Gillin, J. C. State-dependent accessibility of retrieval cues in the retention of a categorized list. *Journal of Verbal Learning and Verbal Behavior*, 1975, **14**, 408-417.

Godden, D. R., & Baddeley, A. D. Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology*, 1975, **66**, 325-331.

Greenspoon, J., & Ranyard, R. Stimulus conditions and retroactive inhibition. *Journal of Experimental Psychology*, 1957, **53**, 55-59.

Kučera, H., & Francis, W. N. *Computational analysis of present-day American English.* Providence, R.I: Brown University Press, 1967.

Mandler, G., & Pearlstone, Z. Free and constrained concept learning and subsequent recall. *Journal of Verbal Learning and Verbal Behavior*, 1966, **5**, 126-131.

Mandler, G., Pearlstone, Z., & Koopmans, H. S. Effects of organization and semantic similarity on recall and recognition. *Journal of Verbal Learning and Verbal Behavior*, 1969, **8**, 410-423.

Pan, S. The influence of context upon learning and recall. *Journal of Experimental Psychology*, 1926, **9**, 468-491.

Puff, C. R. Role of clustering in free recall. *Journal of Experimental Psychology*, 1970, **86**, 384-386.

Smith, S., & Guthrie, E. R. *General psychology in terms of behavior.* New York: Appleton-Century, 1924.

Smith, S. M. Remembering in and out of context. *Journal of Experimental Psychology: Human Learning & Memory*, 1979, **5**, 460-471.

Smith, S. M., Glenberg, A. M., & Bjork, R. A. Environmental context and human memory. *Memory & Cognition*, 1978, **6**, 342-353.

Strand, B. Z. Change of context and retroactive inhibition. *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 202-206.

(Received for publication November 24, 1981;
revision accepted April 28, 1982.)