DOCUMENT RESUME

ED 056 361	24	CG 006 831

AUTHOR	Falkenberg, Philippe R.
TITLE	The Influence of Contextual Change on Remembering in Short-Term Memory. Final Report.
INSTITUTION	Wake Forest Univ., Winston Salem, N.C.
SPONS AGENCY	National Center for Educational Research and Development (DHEW/OE), Washington, D.C.
BUREAU NO	BR-0-C-045
PUB DATE	Aug 71
GRANT	OEG-3-70-0032(509)
NOTE	41p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Experimental Psychology; *Learning; Learning Processes; *Memorizing; *Psychological Studies; *Recall (Psychological); *Retention Studies

ABSTRACT

If the contextual similarity between learning and recall within a single trial in a short-term memory (STM) paradigm is varied, recall varies proportionately. This context effect was demonstrated using variations of the Peterson-Peterson (1959) paradigm for both aurally and visually presented material, verbal and arithmetic context, and within and between Ss designs. Experiments were conducted to discover whether the context effect was due to differential ability to rehearse in the recall interval, differing intertrial intervals, and different amounts of activity on the Ss part during a trial. None of these hypotheses was supported by the data. An experiment was conducted to find out if proactive interference was the cause of the context effect, and this also turned out negative. An explanation of the effect on the basis of stimulus generalization is proposed. (Author)

Final Report

Project No. 0-C-045
Grant No. OEG-3-70-0032 (509)

Philippe R. Falkenberg
Wake Forest University
Winston-Salem, North Carolina 27109

THE INFLUENCE OF CONTEXTUAL CHANGE ON REMEMBERING IN SHORT-TERM MEMORY

August, 1971

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

Office of Education

National Center for Educational Research and Development
(Regional Research Program)

Final Report

Project No. 0-C-045

Grant No. OEG-3-70-0032 (509)

THE INFLUENCE OF CONTEXTUAL CHANGE ON REMEMBERING IN SHORT-TERM MEMORY

Philippe R. Falkenberg

Wake Forest University

Winston-Salem, North Carolina

August, 1971

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
National Center for Educational Research and Development

2

PREFACE

The author would like to thank Dr. Harold Schiffman and Dr. Gregory Lockhead of Duke University for their advice and comment on the first two studies. The assistance of Mr. Frank Wood, Miss Patricia Brown, and Miss Lynn Redfearn in running these experiments was greatly appreciated.

CONTENTS

4

LIST OF TABLES

5

LIST OF FIGURES

Introduction

Changes in context influence learning and retrieval in long term memory (LTM) in humans and animals (Watson, 1907; Carr, 1917; Pan, 1926; Bilodeau & Schlosberg, 1951; Greenspoon & Ranyard, 1957; Thomas and Jones, 1962). Do contextual changes influence learning and retrieval from short-term memory (STM), and if so what mechanisms are operating?

The Peterson and Peterson (1959) paradigm is suitable for the study of contextual influences on STM because: (1) The "busy" task inserted between learning and recall can be used as context; (2) measurements are taken within one trial on single Ss; (3) the context is very different from the learning task and in no logical way can serve to assist learning or recall, on the contrary the context is deliberately chosen to hinder the rehearsal of the learned response (this is important for all too often in studies employing context as independent variables it is not clear where the context leaves off and the to-be-learned S-R association begins); and (4) Ss attention is held to both the learning task and the contextual material.

Nine experiments are reported here; four experiments demonstrate the context effect in STM, namely that retrieval from STM improves as the context for learning is made more similar to the context for retrieval, and five more experiments examine the mechanism by which the context effect might operate.

The first two experiments were conducted before starting this grant, but since the remaining experiments key on them they are included in this report.

Context Effect

Experiment I - Basic Context Effect

In the Peterson-Peterson paradigm Ss learn a consonant trigram, count backwards in the interval between learning and recall, then recall what they can of the trigram. The two criterion events of interest in this paradigm are the learning of the trigram and the recall of the trigram. Events that occur immediately before a criterion event can be construed as part of its context. Note that the context for the recall is very different from the context for learning, because Ss are counting backwards before recall whereas they are not counting before learning. This change of context between learning and recall could account for part of the forgetting in this paradigm. If, as has been found in LTM, context is kept the same for both learning and recall retrieval should be better than when it differs. To test this hypothesis Ss performed the usual Peterson-Peterson task as one condition, and as a second condition, counted backwards before learning from the same number and for the same length of time used in the recall interval.

Method and procedure. Twenty-four Duke undergraduates were paid Ss. Participation in the experiment was a course requirement. Six recall intervals were used: 3, 6, 9, 12, 15, and 18 sec. One condition (PP69) employed the Peterson-Peterson (1959) paradigm unchanged. The other condition (CON) required S to count backwards, from the same number and for the same length of time, immediately before the presentation of the trigram as well as well as between presentation of the trigram and recall. Ss learned consonant trigrams selected from Witmer (1935) with associative values between 4% and 29%. No two successive items contained letters in common, and each trigram occurred equally often over the Ss at each rec i interval. A three-digit number from a table of random numbers was paired with each trigram to provide a starting point for backwards counting. A green light gave a 2-sec. warning before the trial began and remained lit during the trial. A one-second-long red light signalled when E spoke the trigram and later signalled S to say the trigram at the end of each recall interval. Intervals were timed with Hunter timers. A black screen hid E from S. Each S did both conditions, and all recall intervals. In the context condition (CON) another list of trigrams selected from the same source was paired with another set of three-digit numbers, and the same time intervals were used. (Ss counted before learning in this condition, as well as in the recall interval.)

Ss did all trials under one condition before doing the other condition and were separately instructed on the experimental situation before performing each of the two conditions. Ss did two practice trials before each condition. Each S received eight trials for each recall interval for each condition. They had to repeat the trigrams letter by letter after E had spoken them, and they had to say three consonant letters for each recall. The experiment was counterbalanced for lists of trigrams, order in which lists were presented, order of conditions, and the order of recall intervals within each condition.

Results

Table 1 shows the mean proportion of correct responses made by the group for every time interval in each condition. Performance in CON is superior to that in PP69 for every time interval and, except for the 3 second interval, these differences are statistically significant. The performance enhancement of CON to PP69 is roughly equal over the time intervals except for the 3-second interval. The closeness of the results

Table 1

MEAN PROPORTION CORRECT RESPONSES FOR CON, AND PP69

AT EACH RECALL INTERVAL (24 Ss)

	Recall Intervals (sec.)						
	3	6	9	12	15	18	$\bar{X}$
CON	.90	.85	.82	.71	.72	.64	.78
PP69	.89	.71	.65	.58	.64	.52	.66
Difference between CON and PP69	.01	.14	.17	.13	.08	.12	.12
P (one tail)*	.277	.015	.001	.025	.008	.026	<.001

* Determined by a signed ranks test. (Siegel, 1956)

for the 3-second interval for CON and PP69 conditions is a ceiling effect; a major proportion of Ss scored perfectly at this interval (i.e., eight out of eight trigrams), so little difference could be demonstrated.

Experiment I suggests that if the retrieval context is more similar to the learning context, recall is significantly improved. This would mean that part of the forgetting in the standard Peterson-Peterson paradigm is due to contextual change as well as lengthening recall interval. Alternative explanations for this result, such as easier rehearsal under the context condition, will be considered in Experiments V-IX.

Experiments II and III - Quantitative Variation

Can a quantitative manipulation of the contextual similarity between learning and recall produce graded results, and can context be varied independently of the recall interval? Exp. II is described first. Ss added pairs of digits instead of counting backwards as the context task, and stimuli were presented visually on a memory drum instead of aurally.

Method and procedure. Ss were 24 Duke University undergraduates, 11 males, 13 females, all paid. Experimental participation was a class requirement. The 144 consonant trigrams Ss learned had associative values between 4% and 29% (Witmer, 1935). Digits were randomly selected and paired. Every S saw the same set of consonant trigrams and digit pairs. Digits and trigram lists were printed by computer and shown on a Lafayette memory drum. Recall latencies were measured with a Standard timer started automatically by the memory drum paper and stopped by the S releasing a push-button immediately after recall. A black screen hid E from S.

The three recall intervals (5.4, 10.8 and 16.2 seconds) and the four context levels (0, 33, 67, and 100 per cent) combined to make 12 conditions. Each trigram was shown in each condition over Ss. Twelve trials were given for each condition, a total of 144 trials per S. These 144 trials were presented to Ss in the order of a 12 x 12 balanced Latin square (Edwards, 1962) so that each condition followed every other condition once. Different Ss started into the square at different rows so that any practice effect was distributed throughout the square. Ss learned and recalled one consonant trigram in each trial. Subjects added digits saying the totals aloud before learning the trigram as well as before recalling it, to provide similarity of context between learning and recall. To make learning context as similar as possible to recall context, the same additions were used in both cases. Trigrams and digit pairs were presented for .9 sec. In the 100 per cent context level all the digit addition used in the recall interval was performed before learning the trigram. If no digits were added before learning this was the 0% level. A proportion (33% or 67%) of the digits immediately before recall were added before learning to provide intermediate context levels. The actual numbers of digit pairs involved in each of these context levels was determined by how many were used for each recall interval (e.g., a 5.4 sec. recall interval used six digit pairs so a 33% context would use the last two pairs before learning, likewise if 12 digit pairs were used in the recall interval 33% context would use the last four pairs, etc.).

Subjects were seated before the screen directly facing the memory drum, given the push button to hold, and told how to use it. They were instructed that this was an experiment in STM and forgetting, shown how the trigrams and digit pairs would appear in the slot of the memory drum, told to say the trigrams aloud one letter at a time, and to add the digit pairs. Because of the large number of trials Ss were given a one-minute rest every 12 trials, and 72 trials were run on two consecutive days. There was a 9 sec. interval between the presentation of the last digit pair of the recall interval and the beginning of the next trial.

Results

In Table 2 all three measures show that as the degree of contextual similarity between learning and recall increases, retrieval performance increases. By far the most sensitive measure of Ss performance was the error measure (Ss could make 0, 1, 2, or 3 errors on each trial) and it is used throughout the remaining experiments (mean recall errors are pooled over Ss and trials). The interaction between context levels and recall intervals for errors is significant [$F\ (6,138) = 2.5$, $p < .05$], and this implies that the total amount of context is having an effect as well as the relative amount between learning and recall. The largest difference between context levels is between 0% and the 33% context error (planned comparison, $p < .01$), but the error difference between the 33% and the 100% context levels is also significant (planned comparison, $p < .025$).

Subjects take longer to respond when they are wrong than when they are right (wrong $\overline{X}$ = 3.02 secs.; right $\overline{X}$ = 2.27 secs.). The latency differences can be explained on this basis; there were more wrong responses hence longer average latencies for the long recall intervals and the lowest contextual similarities. Latency also increases with number of errors as does the variability of latency.

Subjects' learning (or perception) of the trigrams was not perfect. Of the 10,368 consonant letters Ss were shown, 144 were mislearned as measured by Ss' responses at the time of trigram presentation, and the error total increases as the number of prior additions increases, indicating that Ss probably had more difficulty in breaking set for the longer series of additions, and were caught by surprise when the trigram was shown. (When Ss made learning errors their later recall responses were compared to the wrongly learned response for recall accuracy.) The data shows no apparent relationship between learning errors and the context effect. The number of learning errors is far too small to account for the number and relation of retrieval errors. The context effect occurs at retrieval time. Subjects' number of correct trigrams increased 21 percent and errors decreased 37 percent as contextual similarity increased from 0 to 100 percent.

Table 2

EXPERIMENT II -- MEANS (24 Ss)

	Percent Contextual Similarity			
Measures	0	33	67	100
Proportion Correct Responses	.58	.67	.70	.71**
Recall Errors	.81	.60	.55	.51**
Latencies (sec.)	2.61	2.55	2.48	2.53*

	Recall Intervals (sec.)		
Measures	5.4	10.8	16.2
Proportion Correct Responses	.77	.63	.60**
Recall Errors	.39	.67	.78**
Latencies (sec.)	2.25	2.62	2.76**

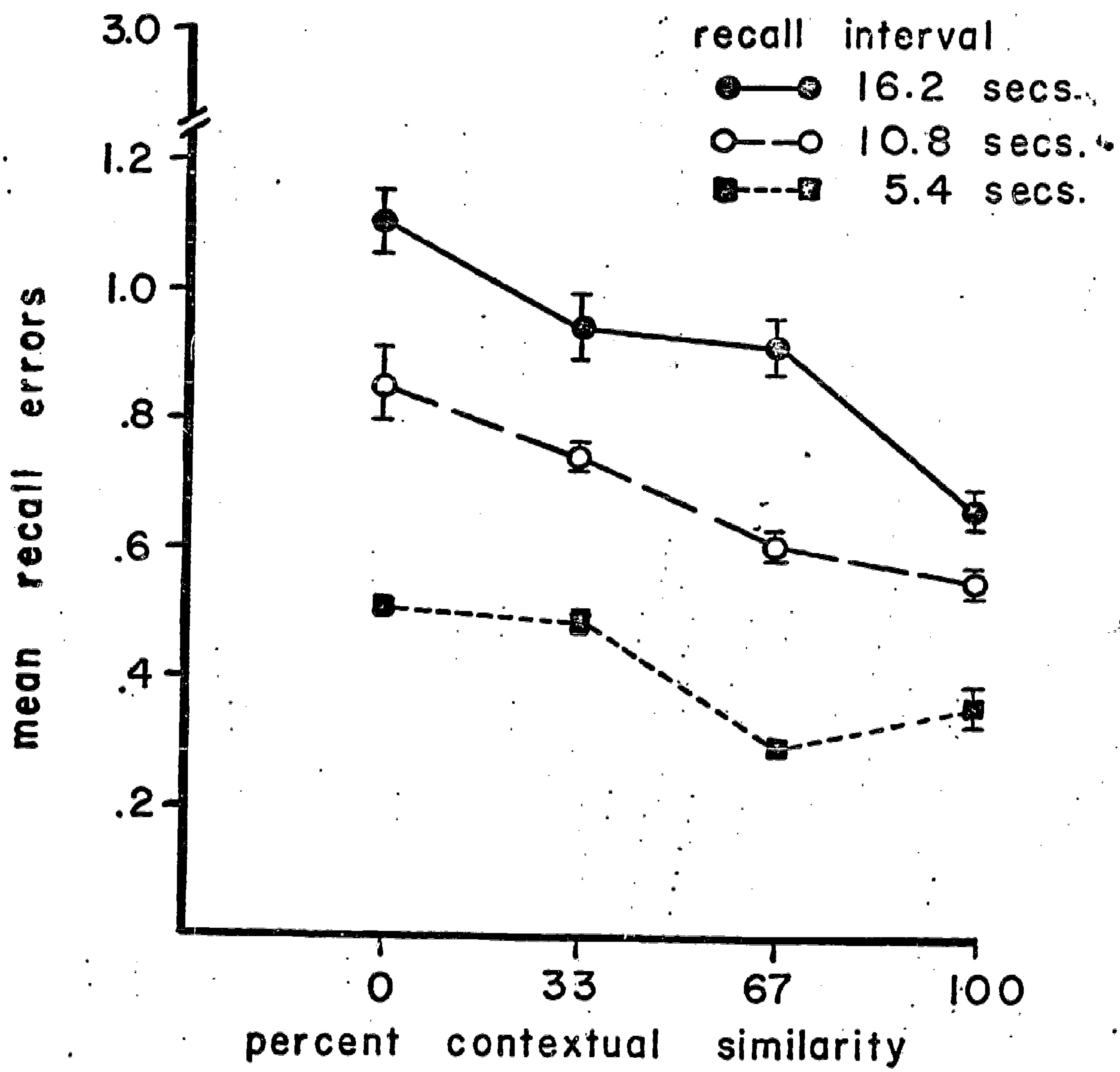
* Means differ at $p < .025$, F test.

** Means differ at $p < .001$, F test.

If contextually similarity between learning and recall is varied quantitatively, Ss ability to recall varies proportionately, and context can be manipulated independently of the recall interval. The usual STM forgetting over increased recall intervals was also found.

Experiment III was a replication of Experiment II with a new experimenter, a better memory drum (Lafayette IBM), a different location, and 12 Ss from Wake Forest University. Latencies were not recorded. With these exceptions everything was the same as in Exp. II. The results shown in Figure 1 are similar to those in Exp. II, and the slope of these

Figure 1. Experiment III--The mean number of recall errors for each recall interval and each level of contextual similarity for 12 Ss.


recall interval
16.2 secs.
10.8 secs.
5.4 secs.
3.0
1.2
1.0
.8
.6
.4
.2
mean recall errors
0
33
67
100
percent contextual similarity

curves are representative of the context effect in all the studies reported here. The recall performance for different levels of contextual similarity are significantly different [$F\ (3,33) = 6.51$, $p < .005$], as is performance for different recall intervals [$F\ (2,22) = 7.38$, $p < .001$].

Data from this experiment was further analyzed showing: (1) when intrusion errors from the previous trial were examined no systematic relationship between them and the context effect was found. The context effect cannot be explained because of proactive interference (PI) from the previous trial. (2) The number of errors was least when the recall interval of the previous trial was the same as the present trial. One might speculate that this "sameness" was providing a form of intertri: ı context effect.

Experiment IV - Independent Groups

The previous three experiments have all employed a within S design. Because of the intertrial effects just mentioned an independent groups experiment was done to ensure that the context effect was not some peculiar artifact of the within Ss design.

Method and procedure. Ss were 64 undergraduate students, 30 females and 34 males. Participation was a class requirement. Materials and apparatus were the same as in Experiment III, except that only 25 trigrams were used.

Ss were randomly assigned to four conditions, 0,,33, 67, and 100 percent context. A 10.8 second recall interval was used. The same list of trigrams was used for each condition, and each S received 25 trials. Ss were given two practice trials and a minute break after every 10 trials.

Results

The mean recall errors for the four conditions were 0%--.82; 33%--.66; 67%--.59; and 100%--.36, and these were significant differences

[F (3,60) = 3.59, p $\angle$ .025]. These results are extremely close to those obtained in Exp. III for the 10.8 sec. recall interval. Orthogonal tests showed significant differences between the 0% and 33% conditions, and the 67% and 100% conditions. The context effect in the within Ss design is not an artifact; because this design is economical of Ss and more sensitive than an independent groups design we continued to use it in the remaining experiments.

Mechanism of the Context Effect

The previous experiments demonstrate that as you change the amount of contextual similarity between learning and recall for individual Ss within a single trial lasting but a few seconds, as similarity increases so does recall. This has been shown to happen for both visually and aurally presented stimuli, and when the context was backwards counting and when context was digit addition. What is the mechanism for the context effect?

The simplest explanation would be stimulus generalization, and the form of the data curves would suggest this hypothesis, but alternate explanations exist and must be considered.

The most likely alternate hypothesis is that the addition of context makes rehearsal in the recall interval easier. Since Ss experienced the same backwards counting or digit additions before learning as they did in the recall interval in 100% context, it could be argued that familiarity with the numbers would make them easier in the recall interval, thereby allowing more rehearsal than in the 0% condition. This was tested experimentally.

A second hypothesis involves intertrial interval. Loess and Waugh (1967) have shown that as the intertrial interval increases in STM recall improves. In the first three experiments reported here this was the case. By adding context before learning the spacing between the trigram on the present trial and the trigram on the previous trial was increased. Possibly this additional time was allowing "better" learning, giving the context effect, so this was tested.

The third hypothesis tested was that differential activity is producing the context effect. In the 100% context condition Ss are busier than in the 0% condition, and possibly this extra activity is "warming them up" so that learning and retrieval are better.

Finally, Kepple and Underwood (1962) propose that the forgetting in this kind of paradigm is caused by proactive interference. The intrusion analysis in Experiment III does not support this explanation for the context effect, although PI clearly occurs; this hypothesis needs further exploration, and is examined in Experiment IX.

Experiment V - Rehearsal

Does familiarity with digit addition allow more rehearsal in the 100% context condition? The number of times Ss saw and added a set of digit pairs within a trial was systematically varied at the same time the percentage of contextual similarity was varied.

Method and procedure. Ss were 36 undergraduate students participating as part of a class requirement. Apparatus, materials, and instructions were the same as Experiment III. Eighty-one trigrams were used. The context levels were 0, 33, and 100 percent. Familiarity with the context was varied by the number of times a particular digit addition set was presented in the recall interval. The digit additions were presented once, twice, or three times in the recall interval; schematically the recall interval task would look like this--once, 123456; twice, 123123; thrice, 121212. The context variable was manipulated as before so that a 33% context level would look like 12 QSX 121212 under a three presentation situation. Each S did all nine conditions of this design, nine trials each; the 81 trials were presented as a balance Latin square as before. Ss rested 1 min. after each nine trials, and all trials were done in one session. The recall interval was 10.8 seconds, and each trigram and digit pair was presented to the S for .9 sec.

Results

Table 3 gives the mean number of errors per trial per S for the three context levels and the three levels of digit pair presentation. The context effect is still present, but the number of times a S experienced the digit additions made no difference in his error scores.

Table 3

MEANS - EXPERIMENT V (36 Ss)

	Contextual Similarity		
	0	33	100
Recall Errors	.66	.55	.52*
	Digit Pair Presentations		
	1	2	3
Recall Errors	.60	.54	.60

* Means differ at $p < .01$, F test.

The difference between the 0% and 33% context levels was statistically significant (planned comparison, $p < .025$). Since the range of digit p[illegible] repetitions went from zero to five over the context conditions (e.g., in the 100%-three presentation condition Ss added the same digit pairs three times before learning and three times before recall) Ss had wide experience with the context familiarity task, yet it made no measurable difference to their scores. Differential rehearsal because of familiarity with the context is not a satisfactory explanation for the context effect.

Experiment VI - Intertrial Interval

In Experiment I-III the interval between trigrams varied, being longer for 100% context than for 0%. The average intertrial interval over levels of contextual similarity was controlled in this experiment by varying the spacing from the end of one trial to the beginning of the next; by doing this the spacing between trigram presentations was held constant for each recall interval.

Method and procedure. The Ss were 24 undergraduates participating as part of a class requirement. Apparatus, materials, instructions, and design were the same as Experiment III, except for the control of intertrial intervals.

Results

Table 4-A-i shows that the context effect is still present even when the intertrial interval is on the average constant, and the usual increase of errors with increasing recall interval is also present (Table 4-A-ii). The magnitude and the form of the context effect data is the same in this experiment as in Experiment II and III.

Because the recall intervals of the previous trial varied, only the average interval between that trial and the present trial is constant. To check whether this averaging was obscuring any result produced by variable intertrial interval the data was re-analyzed using the previous recall interval as one factor and context as another. This permitted us

Table 4

MEANS - EXPERIMENT VI
(24 Ss)

A.-- Average Intertrial Interval Constant

	Contextual Similarity			
	0	33	67	100
(i) Recall Errors	.93	.76	.67	.67*

	Present Recall Interval		
	5.4	10.8	16.2
(ii) Recall Errors	.56	.79	.92*

B.-- Intertrial Interval Varied +

	Contextual Similarity			
	0	33	67	100*
(i) Recall Errors	.96	.80	.68	.71

	Previous Recall Interval		
	5.4	10.8	16.2
(ii) Recall Errors	.83	.80	.74

* Means differ, $p < .01$, F test

+ Based on 132 scores/S. The 1st trial in each of the 12 sets was not included.

to look at the data as if the length of the intertrial interval had been deliberately varied, i.e., the intertrial interval between the present and previous trial would vary according to the length of the previous trials's recall interval over the context levels. Under this analysis the context effect remains (Table 4-B-i) and the intertrial interval effect of Loess and Waugh (1962) appears even though it is not statistically significant (Table 4-B-ii). Loess and Waugh varied their intertrial intervals from 0-60 sec., whereas the intervals in this experiment varied from 19.8 to 41.4 sec. Furthermore, the variation of intertrial interval in this last analysis is completely confounded with the variation of the difficulty of the previous trial. In Experiment III we found that the difficulty of the previous trial did influence performance on the present trial, namely that if the previous recall interval was the same as the present performance on the present recall was best. Analysis on this experiment's data confirm this finding. The effect of intertrial interval on recall in this experiment is probably lessened by this confounding and the difference in the range of intertrial intervals used. The significant finding of this study is that the context effect is not appreciably influenced by intertrial interval.

Experiment VII - Word Context

Before the influence of activity on the context effect can be examined a completely different kind of context is needed. So far context has been an arithmetic task of some kind. Can the context effect be demonstrated using a non-arithmetic contextual task? To answer this question a new context task using words instead of arithmetic was devised.

Method and procedure. *S*s were 18 undergraduates participating as part of a class requirement. With the exception of the context material, apparatus and material was the same as Exp. III. Words chosen from Thorndike and Lorge (1944) (occurs at least 1/1,000,000 list) were shown to Ss in the memory drum slot. Ss were asked to categorize each word as having anything to do with animals or not; in the former case Ss responded "Yes," and in the latter case Ss responded "No," (e.g., Ss seeing CAT, FORK, ELM, QSX, CAT, FORK, ELM would respond yes, no, no, QSX, yes, no, no). Three context levels 0, 33, and 100 percent were used, and a digit addition task was included for comparison giving six conditions for the experiment. Each S received six trials for each condition organized in a balanced Latin square. Ss were instructed in the general nature of the experiment and particularly on both tasks. The recall interval was 10.8 sec. Words, digit pairs, and trigrams were presented for .9 sec. Ss rested for 1 min. every 12 trials and all trials were completed in one session.

Results

Both the word categorization task and the digit addition task produce the context effect (Table 5). Ss recall does not differ on the basis of words or digits, therefore the word categorization task is equivalent to the digit addition task in producing response errors. The gradient of context effect scores for the words task is comparable to that found in Exp. II, III, and VI.

Experiment VIII - Trial Activity

Are Ss being "warmed up" by the additional activity of the 100, 67, or 33 percent context levels in this paradigm? If this is the case, equating the amount of activity over context levels should cause the context effect to vanish.

Table 5

MEANS - EXPERIMENT VII - RECALL ERRORS

(18 Ss)

	Contextual Similarity		
	0	33	100*
Words	.96	.74	.57
Digits	.78	.53	.54

*Contextual levels differ, $p < .05$, F test.

Method and procedure. The Ss were 36 undergraduate students participating as a part of a class requirement. The materials, apparatus, and design were the same as Experiment V, except for the context. Digit addition was used as context for the three context levels, 0, 33, 100 percent. To equate the amount of activity a S did in each trial, the spaces before the 0 and 33 percent were filled with another task. The fill used was: word categorization, a different set of digits than the ones used in the recall interval, or blanks (no activity) as had been used in previous experiments. The recall interval was 10.8 sec.

Ss were instructed on the task as in Experiment V with additional instructions concerning the context tasks. Both the word categorization and digit addition tasks were explained, and Ss were warned that a mixture of the two might appear on any one trial. Ss were given two practice trials on each context task.

Results

Table 6 shows that the context effect remains for each kind of fill. The word and blank fill contribute statistically to the overall significance

Table 6

EXPERIMENT VIII - MEAN NUMBER OF RECALL ERRORS (36 Ss)

	Contextual Similarity			
Type of Fill	0	33	100	$\bar{X}$
Words	.55	.34	.34**	.41
Blanks	.46	.39	.31*	.39
Digits	.40	.37	.35	.38
$\bar{X}$	.47	.37	.34***	.39

* Planned comparison, $p < .05$.
** Planned comparison, $p < .01$.
*** Context means differ, $p < .001$, F test.

of the context effect, but the digit fill does not even though it is in the right direction. The means for each type of fill do not differ significantly, nor do the means for the 100% conditions, which were identical for each type of fill. The overall number of errors is considerably descreased from previous experiments, but this is understandable since the addition of fill would make the 0 and 33 percent context levels resemble the 100 percent level more thereby reducing error rate. The steepest gradient between levels of contextual similarity occurs for the word fill. The contrast between the word categorization task and the digit addition task might be enhancing the similarity of the digit addition context making it more salient. Correspondingly, the shallowest gradient occurs for the different digit fill, and this would be expected since the 0 and 33 percent levels with this fill most closely resemble the 100 percent level. The gradient of the blank fill context levels most closely resembles the form of the gradients in previous experiments. These results negate the activity or warm-up hypothesis as an explanation of the context effect.

Experiment IX - Proactive Inhibition*

Is the context effect caused by differential PI? In this experiment the buildup of PI by trials is examined for a 0% condition, and for two 100% conditions, one of which keeps the same context for every trial and the other changes it each trial. If PI is a suitable explanation for the context effect PI should build up trial-by-trial at a different rate for the 0% and 100% conditions producing an interaction. Analysis of the types of errors made should further elucidate the cause of the context effect.

Method and Procedure. Ss were two hundred sixteen undergraduates, half participated as a class requirement, half were paid volunteers.

Six trigrams were selected from the Witmer (1935), from the 8% and 11% association levels, so that they contained no letters in common. They were randomly ordered into a sequence, and this sequence was then counterbalanced so that each trigram would follow every other trigram once and would occur once in each of the six positions: this procedure yielded six sets of these trigrams. Context words were used as in Exp. VII, and they were randomized into six sets of 20 words each. These sets were then randomly ordered and six counterbalanced sequences of the six sets were produced. Combining each of the trigram sequences with each of the context sequences gave thirty six different conditions of trigram and context sequencing.

---

* This experiment was the Master's thesis of Mr. Frank B. Wood at Wake Forest University.

An 18 sec. recall interval was used. The three experimental conditions were: 0% context (Group 1), 100% context with context changing each trial (Group 2), and a 100% context condition where the context was kept identical for a S's six trials (Group 3). The apparatus was the same as in Exp. III.

The three groups were Group 1, 0% context; Group 2, 100% unique context; and Group 3, 100% repeated context. Each of these groups were divided into the thirty-six trigram and context sequencing conditions, giving 108 conditions. Each of two experimenters ran one subject under each of these conditions. One experimenter was experienced in this line of research but naive to the hypothesis being studied; the other experimenter was Mr. Wood.

Each E ran his Ss independently on days which covered approximately the same period of time. Every three Ss were randomly assigned to the three groups. Ss were instructed about the context tasks and learning trigrams.

Results

No differences were found between the means (.834 and .831 recall errors for the two experimenters so their data were pooled. Significant differences were then found between the groups (context) [F (2,108) = 12.28, $p < .001$], and the trials [F (5,540) = 21.06, $p < .001$], further substantiating the context effect and demonstrating the buildup of errors over trials which is interpreted as increased PI. (Figure 2) Planned comparisons show that the 0% (Gp. 1) and unique 100% context (Gp. 2) were significantly different, but that the repeated 100% context group (3) did not differ from Gp. 2. Ss performing the Gp. 3 context task reported that they became so familiar with the task that they were

Figure 2. Mean total recall errors for Group 1 - 0% context, Group 2-100% unique contex , Group 3 - 100% repeated context.

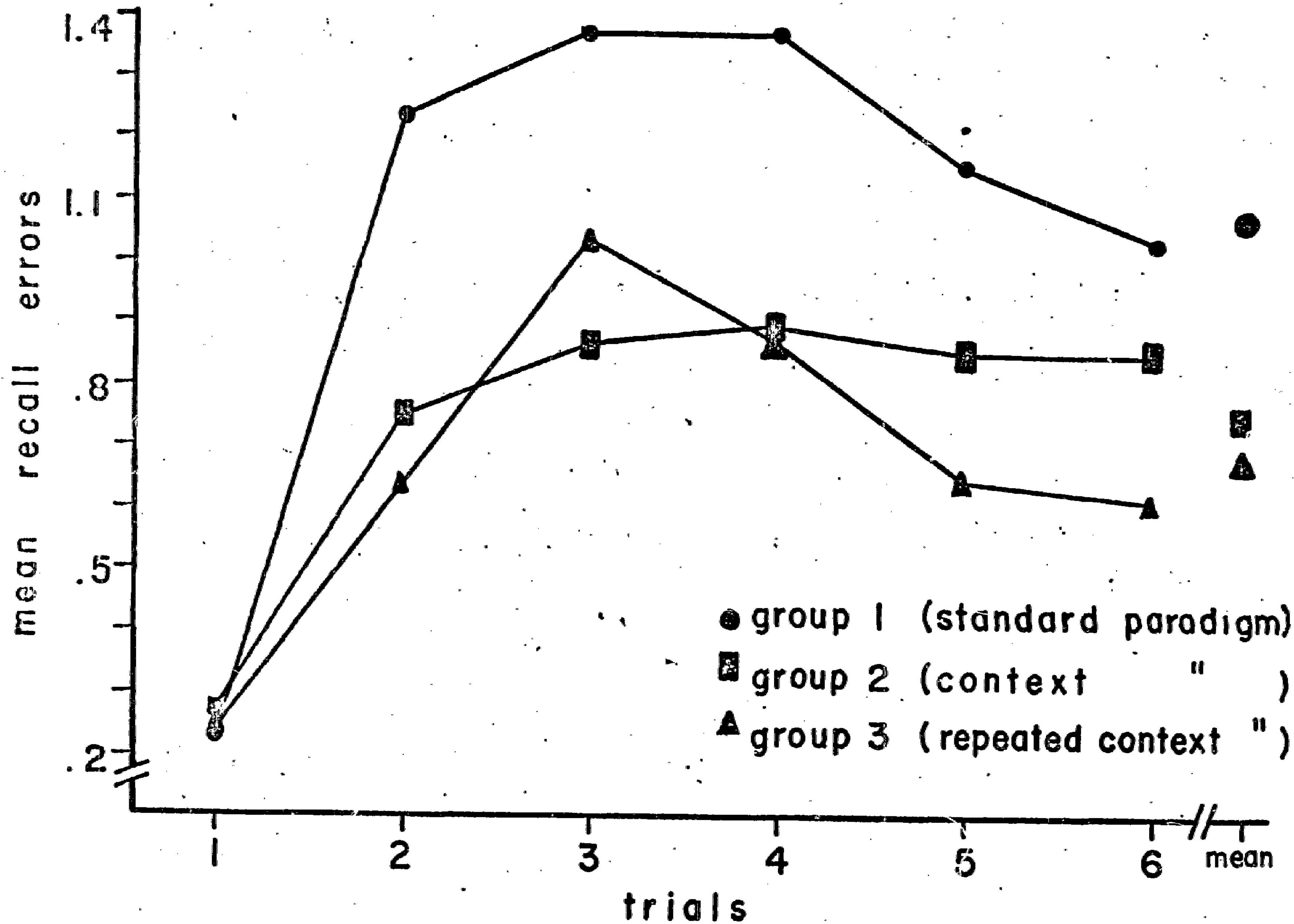

mean recall errors
1.4
1.1
.8
.5
.2
1
2
3
4
5
6
mean
trials
group 1 (standard paradigm)
group 2 (context ")
group 3 (repeated context ")

Figure 3. Mean total and intrusion recall errors, by remoteness of intrusion, for Group 1 (0% context)

Legend:

- t_e total errors from all sources
- t_i total intrusion errors from all degrees of remoteness
- 4 total intrusion errors on trials 5-6 from adjacent trial and two, three, and four trials remote
- 3 total intrusion errors on trials 4-6 from adjacent trial and two and three trials remote
- 2 total intrusion errors on trials 3-6 from adjacent trial and two trials remote
- 1 total intrusion errors on trials 2-6 from adjacent trial

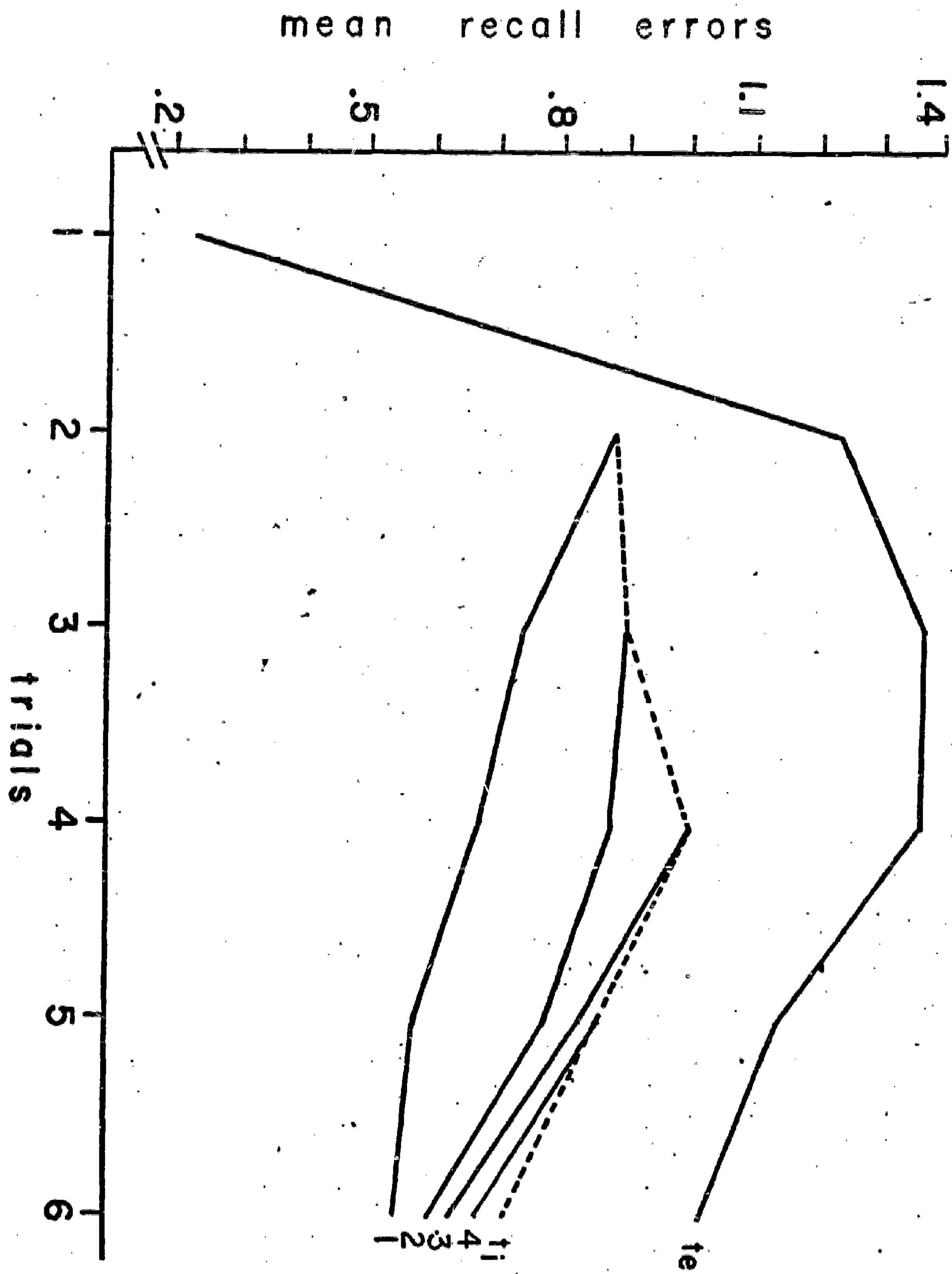

mean recall errors
.2
.5
.8
1.1
1.4
trials
1
2
3
4
5
6
1
2
3
4
ti
te

Figure 4. Mean total and intrusion recall errors, by remoteness of intrusion, for Group 2 (100% unique context)

Legend:

t_e total errors from all sources

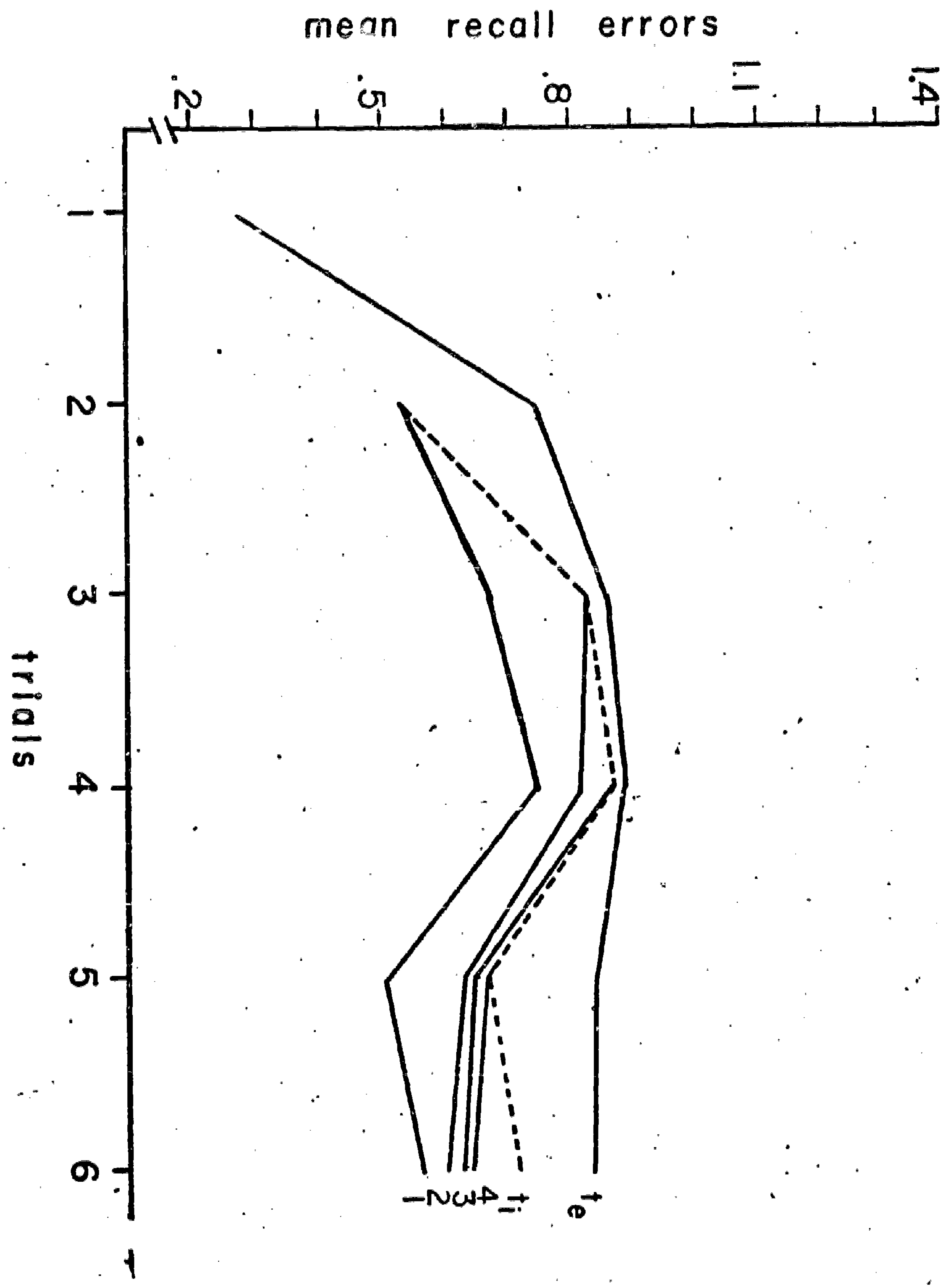
t_i total intrusion errors from all degrees of remoteness

4 total intrusion errors on trials 5-6 from adjacent trial and two, three, and four trials remote

3 total intrusion errors on trials 4-6 from adjacent trial and two and three trials remote

2 total intrusion errors on trials 3-6 from adjacent trial and two trials remote

1 total intrusion errors on trials 2-6 from adjacent trial



mean recall errors
.2
.5
.8
1.1
1.4
trials
1
2
3
4
5
6
1
2
3
4
t_i
t_e

able to rehearse in the recall interval. This rehearsal raises a doubt whether STM alone was being measured, so this group was discounted and will not be considered in further analysis. Scheffe tests indicate that trial 1 differs from trials 2-6, but trials 4-6 do not differ. Most importantly there was no interaction between trials and groups, indicating that the buildup of errors by trial was occurring at the same rate for the context and no context condition.

An analysis of the types of errors made by the context and no context groups shows how they differ. Figures 3 and 4 show that the major distinction between these two groups is the greater number of errors from extra-experimental sources found in the no context group. Three basic types of errors were examined: intra-trial errors (right letters but in the wrong position), intrusion errors from previous trials, and extra experimental errors, which by definition include all errors not in the first two categories. PI is contributing the largest proportion of the errors for both groups (Gp. 1 = .65, Gp. 2 = .81), and intra-trial errors the smallest proportion (Gp. 1 = .10, Gp. 2 = .08). When the intrusion errors are further analyzed by error position it is found that the greatest proportion of intrusion errors are made from the same letter position in a previous trial to the present trial (Gp. 1 = .84, Gp. 2 = .89).

Discussion

Changes in context in a STM paradigm between learning and recall within one trial lasting a few seconds degrade recall performance. In this respect STM resembles LTM. Three possible explanations for this context effect have been explored experimentally, differential rehearsal, effect of intertrial interval, and the effect of amount of activity in each trial. All have been found wanting.

McGeogh (1942) analyzed causes of forgetting in LTM into four mechanisms: stimulus generalization, change of context, interference and change of set. This last mechanism is not pertinent to the phenomenon being discussed here. Stimulus generalization may be the mechanism that accounts for the forgetting due to change of context, depending on how the stimulus is defined.

Of the two forms of interference, proactive (PI) and retroactive (RI), the latter may be dismissed summarily. The S's activity in the recall interval was identical for each level of contextual similarity, so it is difficult to see how differential amounts of RI could have produced the context effect. Furthermore, in the 100% situation the interpolated activity in effect was the second trial for that set of digit additions, whereas the 0% set had only one trial. Since two trials should have produced more learning than one, and RI increases with increased learning of the interfering material, performance should have been worse for the 100% context condition than the 0% condition. The reverse was true.

Although a large amount of PI from the previous trigram was evident, analysis of the intrusions from Exp. III and VI do not show any orderly influence of this interference on the contextual levels. Clearly this is only one source of PI intra-trial, more remote trials or extra-experimental sources may provide the crucial interference.

Experiment IX shows that the context condition has proportionately more intrusion errors than the no context condition, and since intrusions from previous trials is an accepted indicator of PI this weakens the support for PI as the explanation of the context effect. The major distinction between the context and no context groups in that experiment is the excess of extra-experimental errors by the no context Ss. Adding context seems to allow Ss to largely eliminate the extra-experimental source of errors. Admittedly this could be construed as differential PI, but it is a weak argument.

At best the distinction between context and the learned stimulus is arbitrary, and if the context is considered as part of the overall functional stimulus, McGeough's first two causes of forgetting can be considered as just one, stimulus generalization. The gradient of recall errors produced as the amount of contextual similarity between learning and recall is quantitatively decreased produces a stimulus generalization decrement much as is found in LTM. If stimulus generalization is the cause of the context effect, as the total amount of contextual similarity increases the improvement in performance would increase. This occurs in Experiment II, III, and VI as the recall interval increases. Table 7 shows the differences between the average error scores for the 0% and 100%

Table 7

DIFFERENCES BETWEEN 0% AND 100% CONTEXT FOR RECALL ERRORS FOR EACH RECALL INTERVAL

	Recall Interval (sec.)		
	5.4	10.8	16.2
Experiment II	.12	.32	.45*
Experiment III	.16	.31	.44
Experiment VI	.25	.26	.27

* Means differ at $p < .05$, F test.

context levels for each recall interval for Experiment II, III, and VI. There is an increase in performance as the length of the recall interval increases, but only Experiment II shows a statistically significant result. The data suggest that stimulus generalization rather than interference can best explain the context effect in STM.

If the context effect is the result of stimulus generalization the other half of the curve should be produced if the 100% condition is deviated from by adding more context to it rather than subtracting from it as has been done here. Also, discrimination training should sharpen the gradient, and overtraining in the context material should flatten the gradient.

The _Ss_ in the above experiments show remarkably consistent performance in the context effect. Of the 174 _Ss_ run in these experiments in within _S_ designs, 164 showed a difference in error scores between 0% and 100% context, and of these 134 gave fewer errors in the 100% context situation. Similarly 96 of the 143 _Ss_ who were tested with the 33% context level and showed a difference in error scores between 33% and 100% context gave fewer errors at the 100% level. Only 32 of the 58 _Ss_ tested at the 67% level who showed a difference were better at the 100% level, but this difference is not statistically significant although it is in the right direction. The context effect is demonstrable subject by subject, level by level and not just by averaging large groups.

Regardless of the mechanism of the context effect the significance of the above findings is clear. Contextual stimuli do exert a strong influence on people's ability to retrieve from STM if these stimuli are attended to. The orderliness of this influence with quantitative change in contextual stimuli is reassuring to the theorists who must consider context in their theories, for it would appear that contextual stimuli exert their effects in a similar fashion to conditioned stimuli, although not as strongly. The treatment of context by adaptation level theory (Helson, 1964) as a weighted geometric mean just as the conditioned stimulus is, is supported for STM by this data. Similarly stimulus sampling theories of learning (Guthrie, 1935; Estes, 1950) are supported by these findings, since a deliberate change of stimulus sample within a single trial does produce degraded performance as they would predict. These results do not suggest which theory is correct, but they do suggest that theories developed to explain processes in LTM and perception may also be applicable to STM. Finally, the context effect clearly demonstrates the importance of the correct retrieval stimulation in recall for STM.

References

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

Carr, H. Maze studies with the white rat. I. Normal animals. Journal of Animal Behavior, 1917, 7, 259-275.

Edwards, A. L. Experimental design in psychological research. Holt, Rinehart, and Winston: New York, 1962.

Estes, W. K. Toward a statistical theory of learning. Psychological Reviews, 1950, 57, 94-107.

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

Guthrie, E. R. The psychology of learning. Harper, New York, 1935.

Helson, H. Adaptation level theory. An experimental and systematic approach to behavior. Harper and Row, New York, 1964.

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

Loess, H. & Waugh, N C. Short-term memory and inter-trial interval. Journal of Verbal Learning and Verbal Behavior, 1967, 6 (4), 455-460.

McGeough, J. A. The psychology of human learning: An introduction. Longmans-Green: New York, 1942.

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

Peterson, L. R. & Peterson, M. J. Short term retention of individual verbal items. Journal of Experimental Psychology, 1959, 58, 193-198.

Siegel, S. Nonparametric statistics for the behavioral sciences. McGraw-Hill: New York, 1956.

Thomas, D. R. & Jones, . G. Stimulus generalization as a function of the frame of reference. Journal of Experimental Psychology, 1962, 64 (1), 77-80.

Thorndike, E. L. & Lorge, I. The teachers word book of 30,000 words. Bureau of Publications, Teachers College, Columbia University, New York, 1944.

Watson, J. B. Kinesthetic and organic sensations; their role in the reaction of the white rat to the maze. Psychological Monographs, No. 33, 1907.

Winer, B. J. Statistical principle in experimental design. McGraw-Hill: New York, 1962.

Witmer, L. R. The association value of three-place consonant syllables. Journal of Genetic Psychology, 1935, 47, 337-359.