

## Without Surreptitious Rehearsal, Information in Short-Term Memory Decays<sup>1</sup>

JUDITH S. REITMAN

*University of Michigan*

Reitman (1971) found that subjects could retain three words perfectly for 15 sec while detecting tones in noise and supposedly avoiding rehearsal. These results were taken to indicate lack of support for the decay principle of STM. Two studies reported here test two assumptions in the Reitman study: that 100% recall reflects not a ceiling effect but the absence of forgetting, and that lack of disruption of interpolated detection performance indicates lack of rehearsal. Major results indicated that (1) the 1971 study did involve a ceiling effect; (2) tonal detection is measurably disrupted when subjects rehearse; and (3) when subjects detect equally well in the retention interval as in a control interval they forget 33% of what they can recall immediately, and when they detect syllables instead of tones, they forget about 44% more. There is clear evidence for both decay and simple interference in STM.

A classic issue in the area of human short-term memory (STM) centers on deciding which mechanisms best describe the rapid forgetting that occurs. Current theories rely on one of two basic mechanisms: loss of information due to time without rehearsal, called decay, and loss due to displacement of the contents of a limited capacity store by the entrance of succeeding inputs, called displacement interference. The kind of experiment that would decide the issue would require the subject to try to retain something in STM while time elapsed during which no additional inputs entered his STM and he did not rehearse. If the displacement interference mechanism is the only mechanism operative, then the subject in this situation should remember

perfectly. If decay is operative, then the unrehearsed memory trace will have faded and recall should be less than perfect.

Setting actual conditions to fit this ideal experiment are difficult. Subjects typically rehearse when they are not kept busy during the retention interval, and tasks that keep subjects busy, such as counting backward by threes or doing arithmetic, involve successive inputs to STM. Recently, however, a new technique has been proposed which closely approximates these special conditions (Reitman, 1971). In this experimental situation, the subject reads aloud three words, engages in a difficult signal detection task, then recalls what he can of the words. The signal detection task is intended to distract the subject's attention from rehearsal while requiring no new inputs to STM. Check is made against the possibility of the subject's surreptitious rehearsal by testing his detection performance in the retention interval against that in a control interval.

The results of the Reitman study indicated that (1) subjects neither reported nor evidenced rehearsal, and (2) subjects could

<sup>1</sup> This research was supported by NIMH Research Grant No. MH19500. I acknowledge with appreciation the cheerful and able research assistance of Robin Dibner, Barbara Higman, Alan Lifson, and Julia Rosenblum, and the careful critical readings of an early draft by Professors David Kranz and Edwin Martin. Requests for reprints should be sent to Judith S. Reitman, The University of Michigan, Mental Health Research Institute, 205 North Washtenaw Place, Ann Arbor, Michigan, 48104.

Copyright © 1974 by Academic Press, Inc.  
All rights of reproduction in any form reserved.  
Printed in Great Britain

remember virtually all three words after 15 sec of tonal detection, in contrast to remembering only 75% after the same amount of time spent in a parallel task consisting of detecting the syllable *toh* in a mixed series of *dohs* and *tohs*. Atkinson and Shiffrin (1971) and Shiffrin (1973) have replicated these results and extended the time for which there is no forgetting to 40 sec. The results of these two studies indicated little support for the theoretical decay mechanism. Forgetting in STM seemed to be the result of interfering inputs to a limited capacity store, not the passage of time.

Because of the wide implications of the results, the underlying assumptions and details of the experimental technique must be carefully examined. There exists, first, the possibility that the results interpreted as evidencing the absence of forgetting were due to a ceiling effect. The subjects remembered perfectly the three words in the Reitman study and the five consonants in the Atkinson and Shiffrin study. This is not clear evidence that there was no forgetting. It may reflect the fact that the forgetting that did occur was not sufficient to cause subjects to respond incorrectly. Subjects may recall three words or five consonants very easily immediately, have more difficulty after 15 or 40 sec, but still have enough of a trace left to reconstruct an appropriate response.

A second, more central assumption on which the conclusions were based involves the supposed impossibility of concurrent rehearsal and good detection performance. Subjects in the Reitman study were assumed to have avoided rehearsal because they performed a detection task as well in the retention interval as in a control interval. Perhaps subjects can rehearse and detect simultaneously without measurable decrement in performance.

Several recent experiments support the assumption that subjects cannot concurrently rehearse and detect signals. Johnston, Greenberg, Fisher, and Martin (1970), for example, showed that performance on a tracking task in

the retention interval was an inverse function of the subject's memory load, presumably an inverse function of the amount of rehearsal the subject engaged in. And, similarly, Shulman and Greenberg (1971) showed that as list length increased, performance on a perception task in the retention interval decreased in accuracy, and reaction times (RTs) increased. Though generalization of these results to the signal detection task makes the assumption reasonable, direct confirmation is necessary.

Also, in the Reitman study, a two-step procedure was used to determine the equivalence of detection performance in the control and retention intervals, a procedure that is questionably sensitive. In the first step of the analysis, the group's average  $d'$  and RT scores for the experimental condition were compared to those in the control condition and found not to differ significantly. In the second step, when each subject's difference between control and experimental  $d'$  or RT was compared to the differences all subjects produced, none was found to be a statistical outlier (David, Hartley, & Pearson, 1954). That is, no subject's difference was sufficiently unlike that of the other subjects to assume his data came from some other distribution, presumably that of rehearsers.

This outlier test is valid only if the majority of the subjects' performance can be assumed to reflect the avoidance of rehearsal. If more than a few subjects are outliers, or in this case surreptitious rehearsers, their high difference scores will add to the base variance against which these individual scores are tested and none will be detected as statistically deviant. More sensitive, within-subject tests are necessary to validate the claim that subjects are performing as well in the retention interval as in the detection control, and, consequently, avoiding rehearsal.

Two experiments are reported here which are intended to (1) test the no-forgetting results of Reitman, Atkinson and Shiffrin, and Shiffrin for a ceiling effect, (2) test

whether detection performance changes when subjects rehearse, (3) test the two-step analysis for sensitivity in designating rehearsers, and (4) re-examine the question of retention while time elapses without interference or rehearsal.

Experiment I is designed to allow comparison of tonal detection performance with and without rehearsal; Experiment II replicates the 1971 study using both tonal and syllabic detection tasks. Both studies adjust the memory load to correct for the potential ceiling effect. Subjects were presented with five words instead of three and asked to recall immediately on some trials and after 15 sec of detection on others. In Experiment I, subjects performed the tonal detection task in three conditions: in the retention interval of trials in which they were instructed to try to avoid rehearsal, in retention intervals when they were instructed to try to surreptitiously rehearse, and in control intervals with no retention requirements. In Experiment II, subjects were instructed to avoid rehearsal in conditions with the tonal or syllabic detection tasks interpolated in the retention interval and in corresponding control intervals.

## METHOD

In both Experiments I and II, the subject sat in an experimental room by himself, seated at a teletype, three feet in front of an oscilloscope on which the words were presented. He wore Koss Pro-4A earphones throughout the session.

In each trial of both Experiments I and II, five single syllable nouns were presented to and read aloud by the subject, which he then tried to recall either immediately or after 15 sec of performing one of two detection tasks. In what follows, the details of the retention task are described first, then details of the detection tasks, and then general conditions. Descriptions of procedural details specific to each experiment follow.

### *The Short-term Memory Task*

Five common English four-letter nouns of one syllable appeared simultaneously on the oscilloscope for 2 sec. The subject read the words aloud. The beginning and end of the 0 or 15 sec retention interval were denoted by inverted and upright Ts respectively, each presented for 1/2 sec. A row of question marks appeared after the terminal T to signal the subject to recall as many of the five words as he could and type them in order on the teletype. After being allowed 15 sec to type his response, the keyboard locked for a 7 or 22 sec rest interval preceding the beginning of the next trial. Each trial in a block of trials began 40 sec after the beginning of the previous trial, regardless of the length of the retention interval.

### *The Signal Detection Tasks*

Two detection tasks were used, parallel in timing characteristics but differing in the stimuli to be detected.

One was a classic tonal detection task; the other involved detecting the difference between the syllables *doh* and *toh*. During each 15 sec interval, a signal occurred  $n$  times,  $n$  ranging from 0–14. The distribution of the number of signals in a trial was binomial, such that the probability of a signal in the next single sec interval was .5. Whenever the subject heard a signal, he was to press an RT key with his right index finger.

The tonal task consisted of detecting the presentation of a pure tone in a background of white noise. The tone was a 100 msec 1000 Hz square wave; a wide-band white-noise generator produced the background noise. The signal-to-noise ratio was varied for each subject according to a pre-experimental performance criterion of 50% hits.

The syllabic detection task required the subjects to press the RT key each time he heard the syllable *toh* in a mixed series of *dohs* and *tohs*. The syllables were artificially produced sounds made by a synthesizer at

Haskins Laboratory. They were presented at a rate of 1 per sec and lasted 350 msec. They were presented in a background of white noise, the intensity of which was the same as in the tonal detection task. The syllables were equally intense at a level which varied for each subject according to a pre-experimental criterion of 50% hits.

### *Conditions*

Each detection task was performed both as a filler task and by itself in a control condition. In the control conditions, the subject first read five words aloud, then turned his attention to the detection task, knowing in advance that he was not required to remember the words. After the terminal T denoting the end of the detection interval, no question marks appeared. The subject merely awaited the beginning of the next trial.

In the experimental conditions, the subject knew in advance that he was expected to try to recall the five words after the detection interval. After the terminal T of the detection interval, he saw the question marks on the screen and attempted to type as many of the five words as he could remember. Half of the trials tested recall after 15 sec of detection; half tested immediate recall. In the later case, the two Ts appeared in immediate succession, followed by the row of question marks.

Trials were presented in blocks of 12, two blocks devoted to the experimental condition, and one to the control condition. In the experimental condition, trials with 0 sec delays were randomly mixed with those with 15 sec delays, with the constraint that in each block of 12 trials there were six of each. The control trials were run in a single block, since the subject had to know in advance on which trials he was not required to retain the words.

Before the experimental session, each subject experienced a series of trials which allowed the experimenter to set the appropriate signal-to-noise ratios, equating the difficulty of the two detection tasks where appropriate, and making the tasks equally difficult for all

subjects. The subject was then trained with five trials of each detection control condition and five trials of the experimental condition with tonal detection. After the session, the subject was interviewed about his strategies in remembering the words and asked to estimate his success in avoiding rehearsal of the words.

### *Details specific to Experiment I*

The first experiment was intended to test the assumption that performance on the tonal detection task suffers when a subject rehearses. Accordingly, subjects experienced three conditions: a control condition measuring normal ability to detect tonal signals, an experimental condition in which they were instructed to try to retain the words without rehearsing while detecting in the retention interval, and a third condition in which they were told to attempt to rehearse covertly while detecting. In this last condition, the subjects were told to share their attention between rehearsal and detection to the best of their ability while trying to keep their detection performance from suffering. These instructions attempted to encourage subjects to be as surreptitious as possible, so that their data would reflect that of subjects who normally try to get away with a little rehearsal when instructions tell them to avoid it.

The rehearsal condition was always run last in the session, in an attempt to avoid teaching the subject too early how to concurrently rehearse and detect. If he were told to be surreptitious early in the session, he might be more inclined to try the same strategy when told not to rehearse. As a result, the order of presentation of the control and experimental "no-rehearsal" conditions was counter-balanced; the experimental "rehearsal" condition was always last. The confounding of the level of practice or fatigue and instructions was deliberate; data analyses take the confounding into account. Words were rearranged into new 5-tuples in each presentation order so that selection of words was randomized over conditions.

The 29 subjects were right-handed undergraduate University of Michigan males who could type, and were paid \$2.00 an hr for their participation. Because a high proportion of subjects in past experiments reported in the post-experimental interview that they rehearsed illegally "once or twice", and this resulted in discarding their entire set of data, a procedural change was instituted with the intent to salvage data and time. All subjects were given a sheet of paper on which they were to indicate trial by trial their success in avoiding rehearsal. In this way, we could use a large portion of each subject's data and discard only that portion reflecting his momentary reported deviation from instructions. The trial format remained the same; the subject indicated his adherence to instructions during the rest interval of each trial.

#### *Details specific to Experiment II*

Experiment II was intended to be a replication of the 1971 Reitman study with a correction for the potential ceiling effect. Subjects were to try to recall five words either immediately or after 15 sec of detection, whereas in the 1971 study they recalled three words after 15 sec of detection only. A minor change from that study was also made in the characteristics of the detection tasks. Here, the average number of signals per 15 sec interval was seven, whereas in the 1971 study it was .60. This increase in the probability of a signal provided additional data concerning subjects hit rates and RTs during the retention interval, enough data to allow reliable comparison of each subject's performance in the retention interval with that in his control interval.

Each subject experienced four conditions, an experimental and control condition for each of the two detection tasks. Each subject saw six blocks of 12 trials each in one of four counterbalanced presentation orders. Each presentation order involved a new arrangement of words into 5-tuples. The 23 subjects were right-handed undergraduate males from The University of Michigan pool of paid

subjects, paid \$2.00 an hour for their participation, and all having a knowledge of typing.

## RESULTS

### *Does Rehearsal Disrupt Detection Performance?*

Subjects' detection performance is analyzed throughout according to both the two-step analysis introduced in the 1971 study and the new seven-measures analysis described below. After the results of each are described, they are compared for their diagnostic correspondence.

The 29 subjects of Experiment I attempted to detect tonal signals in the retention interval while instructed to surreptitiously rehearse. If rehearsal disrupts detection performance, their data should be distinct from that when they detected in a control interval without retention requirements.

*The two-step analysis.* The first step of the analysis tests whether subjects as a group had different mean  $d'$  and RT measures in their control as compared to experimental conditions. These comparisons take into account a practice effect, estimated by averaging the changes subjects as a group display in performing the detection task first and second. In each case the estimated change in  $d'$  or RT due to practice was subtracted from or added to the control  $d'$  or RT to make it comparable in level of practice to the  $d'$  or RT in the experimental condition. Each subject's adjusted control  $d'$  and RT was then subtracted from his experimental  $d'$  and RT and the mean of these differences scores for the group tested against zero.

These first-step tests showed that when subjects were instructed to rehearse, the average  $d'$  dropped from 1.90 to 1.37,  $t(28) = 6.15$ ,  $p < .01$ , and the RTs increased from 490 to 543 msec,  $t(28) = 6.53$ ,  $p < .01$ . The second step of the analysis indicated that none of the subjects deviated sufficiently from the group to be designated an outlier. The two-step

analysis confirms that rehearsal disrupts subjects' detection performance.

*The seven-measures analysis.* For each subject we recorded: (1) the number of words correctly recalled on each trial regardless of their correspondence to the order of presentation, (2)  $d'$  for each trial, (3) the number of responses in a trial, both correct and false alarms, (4) the average RT to the hits for each trial, (5) the distribution of hits over the five 3 sec blocks of a 15 sec interval, summing over all trials in a condition, and (6) the distribution of RTs over the five 3 sec blocks, combining over all trials in a condition. If a subject was sharing his attention between rehearsal and detection performance, these data should show one or more of the following relationships. We analyze the data according to three potential rehearsal strategies, not all necessarily correlated: (1) consistent rehearsal during the retention interval, lowering overall detection performance as compared to that in the control interval, (2) tradeoffs between

detection and retention performance over trials within the experimental condition, and (3) differing patterns of detection in the 15 sec interval of control and experimental conditions reflecting momentary turn of attention from detection to rehearsal.

Summaries of all measures are given in Table 1.

The first strategy is evidenced by one of two measures, called  $\text{Diff}_{d'}$  and  $\text{Diff}_{RT}$ . Both are within-subject  $t$ -tests comparing  $d'$ 's or RTs in the experimental conditions to those in the control conditions. The subjects in Experiment I under instructions to rehearse had an average  $\text{Diff}_{d'}$  score of 2.70 and an average  $\text{Diff}_{RT}$  score of 2.02. Using a 95% confidence level, 20 of the 29 subjects had significant  $\text{Diff}_{d'}$  scores and 14 of the 28 subjects had significant  $\text{Diff}_{RT}$  scores. Thus, the group as a whole can be considered to have shown significant disruption of detection accuracy and RTs when they were told to rehearse. And, 23 of the 29 subjects individually showed

TABLE 1  
SEVEN MEASURES ANALYSIS OF SUBJECTS IN EXPERIMENT I WITH INSTRUCTIONS TO ENGAGE IN  
AND AVOID REHEARSAL

Measures	When told to rehearse		When told to avoid rehearsal	
	Average score	Number of subjects exceeding .95 criterion	Average score	Number of subjects exceeding .95 criterion
$\text{Diff}_{d'}$ (+ = rehearsal)	+2.70	20	+1.39	10
$\text{Diff}_{RT}$ (- = rehearsal)	-2.02	14	+.44	4
$r_{d'/M}$ (- = rehearsal)	+.04	1	-.02	3
$r_{RT/M}$ (+ = rehearsal)	+.01	2	+.07	4
$r_{R/M}$ (- = rehearsal)	-.02	3	-.13	3
Pattern-hits	—	10	—	7
Pattern-RT	—	8	—	5

some evidence of rehearsal from at least one of these two measures.

The subject engaging in the second strategy changed his allocation of effort from trial to trial, sometimes attending to the detection task and other times attempting to rehearse during the retention interval. We expect a tradeoff between detection and retention performance over trials in the experimental condition. The subject would miss signals or react more slowly while retaining well on some trials, hit more signals or react more quickly while retaining little on others. Or, we might expect that when the subject had a lot of free time in a 15 sec interval, time in which he detected no signals, he might have had a tendency to turn his attention briefly to rehearsal. Consequently, a subject might be expected to show greater retention in trials in which he made fewer overall responses. The three measures taken as evidence of this strategy are called  $r_{d'/M}$ ,  $r_{RT/M}$ , and  $r_{R/M}$ , where  $R$  is the number of responses and  $M$  is the number of words recalled. The first and third are evidence of rehearsal only if negative, the second only if positive. Table 1 lists the averaged correlations and the number of subjects exhibiting significant correlations of each type. Altogether, five subjects exhibited detection and retention data indicating this tradeoff strategy of rehearsal.

A subject engaging in the third, most subtle strategy may have altered his detection performance within a single 15 sec interval. He may, for example, have rehearsed a little at the beginning of the interval, then paid attention fully to the detection task. Or, he may have detected as best he could for the first part of the interval, then feeling he was about to lose the memory trace, have taken his attention away from detection for a brief rehearsal. Because a subject may have increased his effort while actually attending to signals, his average performance in the experimental condition may not appear poorer than that in the control condition, but merely exhibit a different pattern. Figure 1

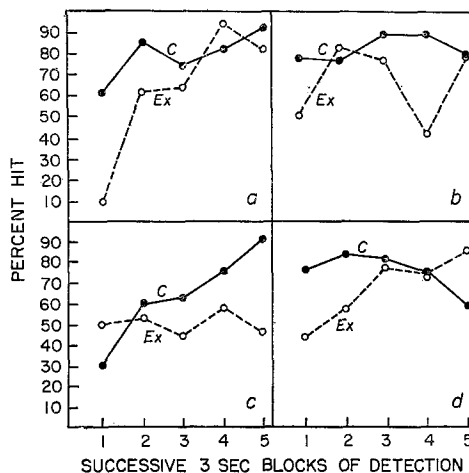


FIG. 1. Experimental and control hit rate patterns inferred to indicate rehearsal.

(a, b, and c) illustrates subjects presumed to have engaged in this type of strategy, rehearsing at the beginning, end, and end of the interval respectively.

Use of such a strategy may be evidenced in a comparison of the hit rates or RTs over five successive 3 sec blocks of detection, with and without rehearsal. The Pattern-hits measure compares the variances of the percent-hit scores for the control and experimental conditions in an  $F$ -ratio, experimental variance divided by the control. The Pattern-RT measure similarly compares RTs in the experimental and control intervals. Using a 95% confidence level, 10 of the 29 subjects had significant Pattern-hit scores; eight had significant Pattern-RT scores. Combining the results of these two measures, 14 subjects individually showed evidence of rehearsal affecting their patterns of detection performance within the 15 sec intervals.

Looking at all seven measures, 25 of the 29 subjects told to rehearse showed at least one piece of evidence that rehearsal disrupted detection performance. What about the remaining four subjects? These four subjects as a group were significantly poorer detectors than the rest of the subjects  $t(7) = 2.65$ ,  $p < .01$ . It may be that these subjects were at a

basement level in performance. Perhaps when they rehearsed, they performed no more randomly than without rehearsal. Signal-to-noise ratios may have inadvertently been set too low for them. One subject, whose detection performance was well within range of the other 25 subjects, had a nonsignificant Pattern-hit measure though his hit-rate patterns for the control and experimental conditions appeared to be distinct (see Figure 1d).

In sum, 25 of the 29 subjects individually showed evidence that rehearsal interfered with their control detecting ability. There is a reasonable chance that the four remaining subjects were special cases, either being at a basement level in detection performance, or exhibiting a rehearsal strategy not detected by one of the seven measures.

Both the two-step analysis and the seven measures showed that rehearsal disrupts detection performance. The seven-measures analysis allows us to specify which individual subjects rehearsed and which of several rehearsal strategies they used. It follows, then, that if a subject is detecting equally well in the retention interval as in the control interval as indicated by the seven measures, it is safe to assume that he is avoiding rehearsal.

#### *Do Subjects Avoid Rehearsal When Told To?*

The 29 subjects in Experiment I also experienced a condition in which they were told to avoid rehearsal while detecting. Comparison of their detection performance in the retention interval with that in the control interval tells us whether subjects can avoid rehearsal when instructed to. As in the previous section, these data are analyzed both with the two-step procedure and with the seven measures introduced in the previous section. Comparison of their results follows.

*The two-step analysis.* The two-step analysis showed that subjects, on the average, detected more poorly when they were told to avoid rehearsal than normal, the average  $d'$  dropping from 1.90 to 1.66,  $t(28) = 3.59$ ,  $p < .01$ .

None of the subjects was designated as an outlier.

*Seven-measures analysis.* As in the case when these subjects were told to rehearse while detecting, we analyze the data according to three types of strategies: overall detection change ( $d'$  and RT) in the retention interval as compared to the control interval, tradeoffs between detection performance and retention over trials within the experimental condition, and differing patterns of detection (both hit rates and RTs) within the interval for the experimental and control conditions. Table 1 lists the average results for the subjects when told to avoid rehearsal.

Looking at all seven measures, subjects as a group had some evidence of disrupted detection when told to avoid rehearsal. The two-step analysis shows that subjects on the average detected more poorly than normal; the seven-measures analysis shows that subjects showed slightly less evidence of rehearsal than when told to rehearse, but still had some evidence of it. Twenty-two of the 29 subjects showed at least one piece of evidence that they were rehearsing in the retention interval when instructed to avoid rehearsal. The two analyses correspond; the seven-measure analysis, however, tells us which subjects in particular were successful in avoiding rehearsal when they were told to.

In this experiment, only six of the 29 subjects showed the kind of differences in detection performance that would indicate strict adherence to the instructions. These six subjects showed at least one piece of evidence of rehearsal when they were told to rehearse, and none when told to avoid it. These six subjects are of critical importance to the issue of assessing retention without rehearsal.

#### *Retention with and without Rehearsal*

Table 2 lists recall performance of the 29 subjects in Experiment I, both when told to rehearse and when told to avoid rehearsal. All recall scores are reported in terms of the percentage of those items a subject could recall



TABLE 2

RETENTION PERFORMANCE: PERCENTAGE OF THAT  
RECALLABLE IMMEDIATELY THAT IS RECALLED AFTER  
15 SEC

	Tonal task	Syllabic task
Experiment I:		
All subjects (N = 29)		
With instructions to re- hearse	98	—
With instructions to avoid rehearsal	71	—
Subjects who followed in- structions (N = 6)		
With instructions to re- hearse	95	—
With instructions to avoid rehearsal	66	—
Experiment II:		
Claimed rehearsers (N = 8)	93	53
Claimed nonrehearsers (N = 15)	85	48
Claimed nonrehearsers with no contrary evi- dence (N = 4)	88	44

immediately that he recalled after 15 sec. The 98% retention with rehearsal is not distinct from 100% recall. When subjects were told to avoid rehearsal, on the average they recalled 71% of what they could recall immediately, a significant loss,  $t(28) = 2.67$ ,  $p < .01$ . Even though most subjects evidenced some rehearsal when told not to rehearse, they forgot significantly more than when they were told to rehearse,  $t(28) = 10.49$ ,  $p < .001$ .

Of critical importance to the conclusions of the studies purporting no evidence in support of decay are the results of the six subjects in Experiment I who showed evidence of rehearsal when told to rehearse and showed no evidence of rehearsal when told to avoid it. When told to rehearse, these six subjects recalled 95% of what they could recall immediately; when told to avoid rehearsal,

they recalled only 66%. This is a significant loss,  $t(5) = 5.59$ ,  $p < .01$ . Using a 95% confidence interval, the true mean lies between 49 and 83%. The retention performance of these six subjects indicates that when there is no evidence that the subjects rehearse, and no interfering inputs presumably enter STM, there is considerable forgetting in 15 sec.

#### *Evidence of Rehearsal in Experiment II*

In Experiment II, eight subjects claimed to have rehearsed,<sup>2</sup> providing performance data that can be compared to that of the 15 subjects in the same experiment who claimed to have avoided rehearsal. These data allow an additional comparison of the sensitivity of the two-step procedure with that of the seven measures in detecting subjects who rehearse.

*The two-step analysis.* The first step of the two-step analysis tests whether the subjects as a group have a different mean  $d'$  or RT in their control as compared to the experimental conditions. Subjects who claimed to have avoided rehearsal showed a significantly poorer experimental RT in the syllabic task condition than in its corresponding control condition, average RTs increasing from 465 to 508 msec,  $t(14) = 2.00$ ,  $p < .05$ . The subjects who claimed to have rehearsed have a significantly worse hit rate in the experimental tonal condition than in the corresponding control condition, average  $d'$  dropping from 1.94 to 1.67  $t(7) = 2.25$ ,  $p < .05$ . All other differences for claimed rehearsers and nonrehearsers were non-significant. The second step of the analysis indicated that there were no outliers in either group.

The preceding analysis categorizes subjects according to their reports of rehearsal and then examines differences in performance of these

<sup>2</sup> Experiment II was run before Experiment I, before the initiation of the procedural change allowing subjects to indicate trial by trial their estimate of success in avoiding rehearsal. The classification here of "claimed rehearser" versus "claimed nonrehearser" comes from the post-experimental interview and relates the claim neither to specific trials nor conditions.

two groups. A stricter evaluation of the two-step procedure involves considering both groups as one and testing the ability of the analysis to identify confessed rehearsers. When the 23 subjects are considered as one group, the mean  $d'$  and RT in the retention intervals were not significantly different from those in the control intervals. Secondly, none of the claimed rehearsers was found to be an outlier. From this we must conclude that the two-step analysis used by Reitman in the 1971 study is not sufficiently sensitive to detect rehearsers.

*Seven-measures analysis.* Table 3 summarizes the average of each of the seven scores for the eight claimed rehearsers and the 15 claimed nonrehearsers for the two task conditions. These results correspond to the two-step analysis only in the overall conclusion that both claimed rehearsers and non-rehearsers show evidence for rehearsal. Details of the kind of evidence they show and in which particular task they show it do not correspond.

Contrary to their various claims, only four of the 23 subjects show no evidence of rehearsal. Table 4 illustrates the corres-

TABLE 3  
SEVEN MEASURES ANALYSIS OF SUBJECTS IN EXPERIMENT II, THOSE WHO CLAIMED REHEARSAL  
AND THOSE WHO CLAIMED NONREHEARSAL

Measures	Claimed rehearsers N = 8		Claimed Nonrehearsers N = 15	
	Average score	Number of subjects exceeding .95 criterion	Average score	Number of subjects exceeding .95 criterion
Tonal Task				
Diff <sub>d'</sub> (+ = rehearsal)	+ .21	1	-.59	1
Diff <sub>RT</sub> (- = rehearsal)	+ .77	1	+ .53	0
$r_{d'/M}$ (- = rehearsal)	-.01	2	-.01	2
$r_{RT/M}$ (+ = rehearsal)	+ .10	0	-.13	0
$r_{R/M}$ (- = rehearsal)	-.17	2	-.13	1
Pattern-hits	—	0	—	3
Pattern-RT	—	1	—	1
Syllabic Task				
Diff <sub>d'</sub>	+2.11	4	+ .68	6
Diff <sub>RT</sub>	-.67	1	-.59	4
$r_{d'/M}$	-.07	0	+ .11	0
$r_{RT/M}$	+ .32	3	+ .11	3
$r_{R/M}$	+ .07	1	+ .10	0
Pattern-hits	—	0	—	1
Pattern-RT	—	1	—	1

TABLE 4

CORRESPONDENCE BETWEEN SUBJECT'S CLAIM OF REHEARSAL AND AT LEAST ONE PIECE OF EVIDENCE OF REHEARSAL

Did subject show evidence of rehearsal?	Did subject claim to have rehearsed?			
	Yes	Yes	No	
		8	11	19
	No	0	4	4
		8	15	23

pondence between the subject's claim of rehearsal and his production of at least one piece of evidence of rehearsal, significant at the 95% level of confidence. According to a Fisher exact probability test, the probability of obtaining this joint distribution given no underlying relationship between claim and evidence is .15. This is not sufficiently low to conclude that there is any correspondence between subjects' reports and evidence of rehearsal. There are too many subjects who claim to have avoided rehearsal who show at least one piece of evidence of rehearsal. In sum, the two procedures match in their conclusions about the claim of rehearsers and nonrehearsers, but the two-step analysis is insufficiently sensitive in detecting liars.

#### *Retention of Claimed Rehearsers and Claimed Non-rehearsers*

Table 2 lists the average retention for subjects who claim to have rehearsed, those who claimed to have avoided rehearsal, and those subjects who both claimed to have avoided rehearsal and showed no evidence to the contrary. All groups except the claimed rehearsers forgot a significant amount when they performed the tonal task in the retention interval,  $t(7) = .76$  for rehearsers,  $t(14) = 4.59$ ,  $p < .01$ , for claimed nonrehearsers,  $t(3) = 8.15$ ,  $p < .01$  for the four subjects. When the interpolated task consisted of syllables rather than tones, all groups forgot significantly more,  $t(7) = 4.72$ ,  $p < 0.1$ ,  $t(14) =$

8.64,  $p < .01$ , and  $t(3) = 16.93$ ,  $p < .01$ , respectively. There seemed to be a general tendency for subjects who claimed rehearsal and subjects who showed rehearsal evidence to recall more, but neither of these differences reached significance in either task condition.

#### DISCUSSION

The conclusions follow the four stated purposes of the work. First, the results of the Reitman (1971) study, 100% recall interpreted as the absence of forgetting, appear to have involved a ceiling effect. The experiments reported here show that when subjects are presented with more than they can recall immediately, they exhibit significant forgetting in 15 sec: 12% when they engage in a tonal detection task, and 56% when they engage in a syllabic detection task.

The two experiments, secondly, support the assumption that when subjects either claim to have rehearsed or are told to surreptitiously rehearse, their detection performance is markedly poorer than control performance. Subjects cannot attend to both rehearsal and detection, confirming the generalization from the work of Johnston *et al.* (1970) and Shulman and Greenberg (1971).

A third purpose was to test the sensitivity of the two-step procedure of the 1971 Reitman study used in identifying subjects who rehearsed while detecting. In Experiment I, comparison of detection performance with instructions either to rehearse or to avoid rehearsal showed a general correspondence between the two-step procedure and the seven measures. In Experiment II, the two-step procedure found the detection performance of subjects who claimed to have rehearsed to be distinct from that of subjects who claimed to have avoided rehearsal. When both groups were combined, however, the procedure did not designate confessed rehearsers as having rehearsed. The seven measures, on the other hand, found evidence for rehearsal not only for all subjects who claimed to have rehearsed,

but also for 11 of the 15 subjects who claimed to have avoided rehearsal. These results suggest that the two-step analysis is inadequate for identifying rehearsers. The seven measures, on the other hand, indicate in detail which subject rehearsed, which of several rehearsal strategies he used, and how strongly he rehearsed, in terms of the magnitude of the disruption.

Finally, of critical importance to the classical issue to which this research is directed is the retention performance of the 10 subjects who followed instructions to avoid rehearsal. These 10 subjects include six from Experiment I who were told to avoid rehearsal and showed no evidence to the contrary, but who did show evidence of rehearsal when told to rehearse, and four subjects from Experiment II who claimed to have avoided rehearsal and showed no significant evidence of it. After 15 sec of tonal detection, subjects in Experiment I recalled 65% of what they could recall immediately; subjects from Experiment II recalled 88%. This is clear evidence for decay. Information in STM is lost when time elapses without interference or rehearsal.

Why are these conclusions inconsistent with those of Reitman (1971), Atkinson and Shiffrin (1971), and Shiffrin (1973)? There are two potential sources of error in the earlier work: ceiling effects and undetected covert rehearsal. The earlier Reitman study potentially suffers from both sources. Subjects remembered three words perfectly, and the tests used to identify covert rehearsers have here been shown to be insufficiently sensitive.

The work reported both in Atkinson and Shiffrin and in Shiffrin, though reporting 100% recall, seemed to avoid the ceiling effect problem. They argued that if there were a ceiling, then a trace weakened by 40 sec of detection would be more disrupted by a small amount of interference than a trace after 8 sec of detection. In their Experiment III, however, the trace after 40 sec was as affected by short

and long periods of disruption as the trace after 8 sec.

Their work, however, does potentially suffer from the problem of covert rehearsal. Their subjects claimed to have avoided rehearsal; as a group their subjects' performance in the retention interval matched that in the control interval; and, the subjects' recalled less than when they were told to rehearse overtly. This argues only that subjects were not rehearsing as much in the avoid-rehearsal condition as when told to rehearse overtly. It does not prove that they avoided rehearsal altogether. As in the Reitman study, subjects may have engaged in various surreptitious rehearsal strategies detectable not on the basis of group performance, but only with sensitive within-subject tests. Many subjects in the present work who were told to avoid rehearsal and claimed to have succeeded in avoiding it had strong behavioral evidence of rehearsal.

The current work additionally allows assessment of the forgetting caused by a small amount of interference, that attributable to the interpolation of two repeated syllables to which the subject must carefully attend. In the condition in Experiment II in which subjects performed the syllabic detection task in the 15 sec retention interval, considerable forgetting took place. The four subjects who showed no evidence of rehearsal forgot 56% of the words they could recall immediately. By subtraction, we can infer that mere interpolation of simple verbal material causes an additional 44% forgetting over the loss from decay in time. Both decay and displacement interference affect forgetting.

Several models of STM have previously posited the existence of both decay and displacement interference. Wickelgren (1970), for example, modeled decay in STM as a mechanism that produces forgetting both as a function of the number of intervening items and the passage of time. Massaro (1970) modeled a decay rate that varied as a function of the similarity of other information to be processed. And, Reitman (1970) modeled the

coexistence of decay and displacement interference in a queueing model of information processing in STM. The experiments reported here confirm the existence of both mechanisms and suggest a procedure for determining independently the size and form of each mechanism's contribution to forgetting in STM.

## REFERENCES

- ATKINSON, R. C., & SHIFFRIN, R. M. Control processes in short-term memory. *Scientific American*, 1971, **224**, 82-90.
- DAVID, H. A., HARTLEY, H. O., & PEARSON, E. S. The distribution of the ratio in a single normal sample, of range to standard deviation. *Biometrika*, 1954, **41**, 482-493.
- JOHNSTON, W. A., GREENBERG, S. N., FISHER, R. P., & MARTIN, D. W. Divided attention: A vehicle for monitoring memory processes. *Journal of Experimental Psychology*, 1970, **83**, 164-171.
- MASSARO, D. W. Perceptual processing and forgetting in memory tasks. *Psychological Review*, 1970, **77**, 557-567.
- REITMAN, J. S. Computer simulation of an information processing model of short-term memory. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970. Pp. 117-148.
- REITMAN, J. S. Mechanisms of forgetting in short-term memory. *Cognitive Psychology*, 1971, **2**, 185-195.
- SHIFFRIN, R. M. Information persistence in short-term memory. *Journal of Experimental Psychology*, 1973, **100**, 39-49.
- SHULMAN, H. G., & GREENBERG, S. N. Perceptual deficit due to division of attention between memory and perception. *Journal of Experimental Psychology*, 1971, **88**, 171-176.
- WICKELGREN, W. A. Multitracé strength theory. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970. Pp. 65-102.

(Received August 29, 1973)