This study examined the effects on long-term retention of variations in intensity and of temporal parameters of arousal following a single learning trial in a paired-associate task. The subjects were 56 female university students. Intensity of arousal was manipulated by using two levels of white noise—75 decibels and 90 decibels sound pressure level—and a condition without white noise. Noise was delivered at three temporal intervals following the learning trial: under three minutes, three to six minutes, and six to nine minutes. The results were analyzed using six linear contrasts. The main effect of level of arousal was not significant. Neither was there any significant interaction between intensity and timing of noise. Effect of temporal variation was examined via two contrasts. It was found that stimulation between three and six minutes after learning was detrimental to retention as compared to stimulation within three minutes and between six and nine minutes. (Author/JM)
Post-Learning Arousal Change and Long-Term Retention

V. K. Kumar
Department of Education
Case Western Reserve University
Cleveland, Ohio 44106

Frank H. Farley
Department of Educational Psychology
University of Wisconsin
Madison, Wisconsin 53706
Abstract

The study examined the effects on long-term retention of variations in intensity and temporal parameters of arousal following a single learning trial in a paired-associate task. The subjects were 56 female university students. Intensity of arousal was manipulated by using two levels of white auditory noise: 75 dB and 90 dB SPL and a condition without white noise. Noise was delivered at three temporal intervals: 0-3 min., 3-6 min., and 6-9 min., following the learning trial. The results were analyzed using six linear contrasts. The main effect of level of arousal was not significant. Neither, there was any significant interaction between intensity and timing of noise. Effect of temporal variation was examined via two contrasts. It was found that stimulation between 3-6 minutes was detrimental to learning as compared to stimulation between 0-3 minutes and 6-9 minutes.
Post-Learning Arousal Change and Long-Term Retention

Studies on arousal using paired associate (PA) learning materials have largely been correlational-relating arousal properties of stimuli to their retention. There have been relatively few studies on studying the effects of experimentally inducing arousal on retention of PAs. Such studies have generally concentrated on arousing prior to learning or during learning or just prior to the retention tests (e.g., Alpern, 1948; Berlyne and Lewis, 1963; Berlyne, Borsa, Hamacher, Koenig and Isolde, 1966; Haveman and Farley, 1969; and Uehling and Sprinkle, 1968). To the authors' knowledge, there has been only one study which investigated the effects of post-learning arousal using PA learning materials (Farley and Lovejoy, 1968). Farley and Lovejoy administered the arousing stimulus following the completion of the last (2nd) trial rather than following each response item as in Berlyne et. al. (1966) study. Berlyne et. al. (1966) designed their study on the notion that the principal role of arousal is in the consolidation of memory trace during a period of perseveration involving reverberatory circuits following the response. Their study, however, did not reveal any significant differences between the different arousal conditions, i.e. (a) arousal during the presentation of the stimulus (4 seconds) alone and stimulus and response terms together (2 seconds), (b) during the interval between items (6 seconds), (c) during the entire period of 12 seconds (the presentation of the stimulus, the stimulus and response terms and the interval). They did, however, obtain a significant difference between the noise and no noise conditions. There seem to be two major weaknesses of Berlyne, et. al. (1966) study: (a) they used 3 trials instead of one and hence the effects of rehearsal
may be confounded with the effects of arousal, and (b) Berlyne and Lewis (1963) in an earlier study had reported that white noise raised one index of arousal, i.e., skin conductance and kept it raised for at least 10-15 minutes. If this is so, then the three white noise conditions should be quite equivalent from the point of view of arousal and hence their recall scores should not be very different, with the small differences attributable to chance. In this respect Farley and Lovejoy's (1968) decision to administer white noise following the completion of the learning trial appears more appropriate. They used six nonsense syllable-familiar word pairs as PA learning task. They delivered white noise (75 dB SPL) to three different groups between 0-3 minutes, 3-6 minutes and 6-9 minutes following the last learning trial at two levels of retention test (12 minutes and 24 hours). A control group at each level of retention received no white noise. Comparing the results of 12 minutes and 24 hours retention tests it was found that the control groups' recall decreased by almost 30%. Recall decreased by 14% and 8% in the case of the 0-3 minute and 6-9 minute groups, respectively, while in the condition 3-6 minutes there was a marked reminiscence effect by almost 20%, quite contrary to their expectation of decreasing recall from condition 0-3 minutes to 6-9 minutes in a linear fashion.

An overview of the literature suggests that there are two major viewpoints concerning arousal and retention. One approach strongly favors a positive relationship between arousal and retention, such that high arousal is detrimental to immediate recall and facilitative of long-term recall. On the other hand, another view strongly supports an inverted "U" relationship between arousal and retention. However, no definite conclusions can be drawn regarding the optimal level of arousal and the
optimum time of arousal from the studies reported to date. As has been seen, most studies with human subjects have been concerned with arousal during learning, with very little attention paid to post-learning arousal manipulation. The study of Farley and Lovejoy (1968) was of the latter type and served as a basis for the present investigation with certain modifications. Farley and Lovejoy (1968) worked with two levels of noise, i.e., 75 dB SPL and a no-noise condition, and provided two training trials on nonsense syllable-word pairs. It is possible that their results were confounded with rehearsal effects. Also, their design allowed inferences about time but not the intensity of arousal.

With these considerations, the present study was designed to examine the influence of experimentally induced arousal after learning, manipulating both the time and intensity of arousal on long-term retention. Only long-term retention was studied, so as to ensure that we were dealing with durable learning effects and not merely transient performance effects (Berlyne, 1967).

Two levels of white noise manipulation (75 dB and 90 dB SPL (Re: 0.0002 dyne/cm²) were chosen with a no-noise condition constituting the control group. The latter group did not represent entirely a "no-noise" condition, it essentially means that no white noise (NWN) was administered to this group. This also constituted the low-arousal condition in the study. The 75 dB SPL was chosen so as to obtain a moderate level of arousal (Chase and Graham, 1967) and also due to the fact that 75 dB SPL intensity has been used in earlier studies of Berlyne et al. (1965, 1966), Haveman and Farley (1969), and Farley and Lovejoy (1968). The 90 dB SPL noise level was chosen so as to obtain a high-arousal condition. Gibson and Hall (1966) reported that stimuli between 85 dB and 100 dB were
judged reliably as being distractive and noxious and presumably resulting in a higher level of physiological activation. To keep noise conditions below damaging limits, no stimulus higher than 90 dB SPL was chosen (Harris, 1957). Yet another consideration in the study was the form of noise. Earlier studies had used a continuous form of noise. Since the noise was to be delivered for a total of 3 minutes in the present study, the duration employed by Farley and Lovejoy (1968), it was considered that a continuous form may result in fatigue or habituation effects. Therefore the noise was shaped and pulsed through an electronic switch before delivery to the subjects' earphones.

There were three temporal intervals of post-learning arousal manipulation chosen in this study -- 0-3 minutes, 3-6 minutes, and 6-9 minutes; these parameters were based on the Farley and Lovejoy (1968) study. The design of the study is presented in Figure 1.

(Figure 1 about here)

In brief, the experiment was designed to study:

1) The effects of three levels of arousal (no noise, 75 dB and 90 dB SPL) induced by white noise after a PA learning trial on long-term retention (24 hours). On the basis of earlier studies, it was expected that the noise conditions would facilitate long-term retention as compared to the no-noise condition.

2) The effects on long-term retention of inducing arousal by white noise at three temporal intervals, 0-3 minutes, 3-6 minutes, and 6-9 minutes, following a learning trial.

3) If there is any interaction between the level of arousal (noise intensity) and timing of arousal in influencing long-term retention (24 hours).
Method

Learning Task

Six PAs used by Kleinsmith and Kaplan (1964), each consisting of a stimulus term and single-digit response term were presented during the learning trial. The stimulus words alone were used during the recall period. The stimulus words were CVC nonsense syllables of zero percent association value (to obtain "random" arousal effects); namely CEF, QAP, TOV, JEX, LAJ, and DAX. The response terms were single digits from 2 to 7 respectively. The stimuli and the stimulus-response pairs were presented on 2" x 2" (50.8 mm x 50.8 mm) slides. Two 2" x 2" (50.8 mm x 50.8 mm) color slides each containing five colored spots arranged horizontally in two rows (red, green, orange, brown, yellow, and blue were randomly used on these slides), were inserted between the PAs to separate the arousal effects of one stimulus from the next (Kleinsmith and Kaplan, 1964).

Interpolated Task

Two mazes were used as the interpolated task following the completion of the learning trial. Some considerations in selecting the mazes as a task were

(1) the task shouldn't be uninteresting;

(2) should be simple and involve a minimum of thinking or mental activity, so that there is minimum of interference from such processes during the consolidation period (Posner and Rossman, 1965) and Weiner (1967) have demonstrated that retention is greatly reduced as the difficulty of an interpolated task increases;
(3) should not be too arousing or tiring or related to the learning task.

Farley and Lovejoy (1968) had used 17 random polygons and had instructed subjects to rate them on a 1 to 7 scale on dimensions of interestingness, complexity, pleasingness, dullness, unusualness, and dislike. Mazes were preferred to the polygons because it was suspected that they might induce arousal and additionally that rating on a 1 to 7 scale would be working with digits as used in the learning task (Farley and Lovejoy did not employ digits in their learning task). Another consideration was that the task should be experimenter paced and not subject paced because subject pacing may involve some subjects working very rapidly and others very slowly, thus inducing differential arousal effects depending upon what they understood the purpose of the task to be, as well as reflecting such individual difference factors as personal tempo (Rethlingshafer, 1963). Hence, each subject worked on each maze for 7.5 seconds. The timing was determined from a pilot study with three subjects who considered 5 seconds to be too short and 10 seconds too long to satisfactorily complete a maze.

The time interval for working on the mazes was controlled by a programmed light flash (1.5v., .075 A) through the use of a Cousino Synco-Repeater Model SR-7341. The light bulb was fixed at a convenient distance and height for the subject, and was covered by a cap during the learning trial. The Cousino was kept outside the testing chamber to minimize noise effects.

Testing Conditions

The slides for learning and recall were projected from outside an Industrial Acoustics Company Model 120 2A acoustic chamber to a projector.
screen covering one of the chamber's windows, thus effecting back projection presentation. The chamber had a measured ambient noise level of 56 dB SPL assessed with Bruel and Kjaer equipment described below, with two persons in the booth. The chamber was lighted with a 60-watt incandescent bulb turned away from the subject's face, thus providing one foot candle luminance at subject's face as measured by a General Electric DP-9 light-meter.

Arousal Equipment

White noise was generated by a Grason-Stadler Model 901B white noise generator. The signal from the generator was shaped and pulsed (125 msecs. on and 125 msecs. off, with a rise and decay time of 25 msecs.) by a Grason-Stadler Model 829E electronic switch, transmitted through a General Radio Model 1450 alternator and finally delivered to the subject via TDH 39 earphones mounted in MX 41/A-R cushions. Pulsed noise was preferred to a continuous noise condition to avoid fatigue and habituating effects.

Prior to running the experiment the acoustic output of the system was calibrated to 90 dB and 75 dB SPL with Bruel and Kjaer apparatus consisting of the following components: artificial ear H152, 6cc. coupler NBS - 9A, condensor microphone 4132, cathode follower 2613, and audio-frequency spectrometer 2112. The output of the system was checked by monitoring the voltage across the terminals of the earphones and the values of tolerance were found to be between ± 0.5 db. All the apparatus were outside the sound proof booth and controlled with a switch inside.

Audiometric Screening Equipment

Prior to participation in the experiment, each subject's hearing was screened bilaterally at 15 dB - HL (ISO-1964) with a Beltone Model 15c audiometer on the following frequencies: 250, 500, 1000, 2000, 4000, and
6000 Hz (0.25, 0.5, 1, 2, 4, 6 Kc/s). This was done to insure that only subjects with normal hearing were included in the experiment. An intercom was used at the time of screening since the experimenter was outside the testing booth.

**Miscellaneous Equipment**

The following equipment was used: GSR transducer, connecting cord, and a projection screen.

**Design**

Subjects were randomly assigned to one of the seven conditions with eight subjects per cell (See Figure 1.). To correct for serial order effects, six different training lists were generated so that each of the six PAs appeared once in each ordinal position in the list (Fisher and Yates, 1938). Since there were eight subjects, two of the lists were randomly picked from out of the six and were used twice. Subjects were randomly assigned to lists and conditions using the block randomization method. The lists were similarly assigned for the recall test but with a restriction that the same subject did not get the same list order for the learning and the test trials.

**Subjects**

Subjects were 56 female students drawn from different undergraduate courses in Educational Psychology and Art Education. Students taking Educational Psychology were given one hour credit for their participation. Nevertheless, their participation was voluntary. Students who were involved in other learning experiments were not included in the study. The participation of Art Education students was purely voluntary. Students had to be taken from a department other than Educational Psychology primarily because a number of the latter students had already been involved in
other learning experiments prior to the present study. Only those students who had had no ear or hearing problems were asked to participate.

Procedure

The windows of the test chamber were covered to eliminate any visual distractions. Subjects were seated comfortably in a padded chair located in the booth.

Subjects were first told that the major purpose of the experiment was to take a series of physiological measures while they performed various tasks. These instructions were given to divert the attention of the subject from the main (learning) task, thus preventing rehearsal and to disguise the nature of the task to be given 24 hours later. (The latter was ostensibly to give an estimate of any physiological changes over 24 hours.) Subjects were then screened for normal hearing. Subjects indicated over the intercom whenever they heard a tone. The criterion of rejection for abnormal hearing was set at a failure to detect tones at two or more frequencies (Newby, 1964). Following audiometric screening, the GSR transducer was attached to the subject's nonpreferred hand with a back-of-hand to palmar placement. She was informed that GSR, a harmless, simple physiological measure, would be automatically recorded outside the chamber. No skin preparation was undertaken, nor were the sponge electrodes impregnated with electrode jelly, as real GSR recordings were not in fact taken.

The learning phase involved presentation of the PAs; during the learning trial, each stimulus term appeared alone for 5 seconds, followed by the stimulus and response term for another 5 seconds. The PA slides were followed by two colored slides for 5 seconds each. Subjects were, however, given a 10-second familiarization period with the color slides
prior to the learning trial in order to reduce any arousal effects that might be attributable to the presentation of these slides during learning. Also during this period, the experimenter verbally labelled the colors for the subject and ascertained that she experienced no confusion or difficulty in discriminating them. Two color slides then preceded the first PA so that the subject could "settle down" before the PAs were presented. Subjects were instructed to "concentrate carefully on both the colors and the nonsense syllable number-pairs and to call them out loud," but to avoid rehearsal they were not specifically told that they would be tested for recall.

After the learning trial, all subjects worked on a set of mazes for 12 minutes. Instructions for the mazes were given immediately after the instructions for the learning trial. Subjects were also familiarized with both the mazes and the light flash signal at the time of giving instructions, for the same reason mentioned earlier, and were instructed to start working on the mazes immediately after the learning trial.

During the consolidation period, the experimental subjects (six conditions, n = 8), received white noise through earphones. There were two levels of noise intensity--75 dB and 90 dB--delivered during 0-3 minutes, 3-6 minutes and 6-9 minutes following a single learning trial. Subjects were told a few seconds before the presentation of white noise that they would hear a sound in their ears which would not hurt them. Subjects wore earphones only during the noise interval to avoid anxiety or arousal from further anticipation of receiving noise. Subjects continued to work on the mazes throughout the 12-minute period, including the period of noise presentation. All subjects were required to work on the mazes for 12 minutes for two main reasons. First, the condition which received
noise during 6-9 minutes would receive at least 3 minutes of maze performance after the induction of arousal so as to avoid any special interference immediately after the presentation of the arousal stimulus. Second, all subjects worked on the mazes for the same amount of time in order to insure constancy among conditions (see Figure 1). The control group of the no white noise condition did not wear earphones at any time and simply worked on the mazes for 12 minutes. This was done to avoid any anxiety due to wearing earphones since noise was not to be presented and it would have been difficult to convince subjects to wear them without reason. Subjects were then requested to come next day at the same time for another set of physiological measures.

During the recall session, subjects were again instructed that the purpose of the experiment was to take physiological measures to insure constancy of conditions. The GSR transducers were again used. Stimulus words alone were presented for 5 seconds each and subjects were instructed to recall the correct number and to guess if uncertain. The correct numbers were not repeated. Color slides were used as before as an interpolated task. Subjects were then once again screened for hearing, to emphasize the physiological nature of the experiment, and requested not to mention the experiment to their fellow students.

Results

In Table I is located the mean number of correct responses in the 24 hour retention test for the different conditions of intensity and timing of arousal. The figures in parenthesis indicate the mean percentages.

(Table I about here)
Since the design did not completely conform to a factorial analysis of variance design, the results were analyzed using linear contrasts. Contrasts that appeared meaningful were designed considering the means. Six contrasts orthogonal to each other and which seemed to account for information of interest were tested by computing $F$ ratios (Hays, 1963). The contrasts and their $F$ ratios are presented in Table II.

(Table II about here)

The mean square error (0.5917) for the analysis in Table II was derived from all the cells and each contrast was tested using 1 and 49 degrees of freedom.

The analysis showed no significant differences between the noise and the no noise conditions (contrast 1). The main effect of noise tested by contrast 2 was not significant. The main effect of temporal variation ($df = 2$) was examined via two 1 df contrasts. Contrast 3 revealed a significantly inferior performance in the 3-6 minutes condition as compared to the combined average performance of the 0-3 minutes and 6-9 minutes conditions. Contrast 4 indicated no significant difference between 0-3 minutes and 6-9 minutes conditions. Finally, contrasts 5 and 6 were designed to test the interaction effects between noise intensity and timing of noise. Both the contrasts were not significant; contrast 5 compared the difference between 75 dB and 90 dB conditions of 3-6 minutes with the average difference between the same noise conditions of 0-3 minutes and 6-9 minutes conditions. Contrast 6 compared the difference between 75 dB and 90 dB noise conditions of 0-3 minutes with the same difference of the 6-9 minutes condition.

Since subjects were drawn from two departments (Educational Psychology and Art Departments) and students from the former got credits for
participation and the latter volunteered. Possible confounding of results may have occurred although subjects were randomly assigned to the treatments. This was analyzed by testing the difference between the means of the two groups. The means of Educational Psychology students ($N = 24$) and Art Department students ($N = 32$), were 0.8750 and 1.0625 respectively. The "t" value obtained ($t = 0.8704$) for uncorrelated groups was not significant ($p > .05$, two tailed).

**Serial Effects**

Analysis was undertaken to find out if there were any effects on recall due to serial position during learning. Frequency of correct responses were tabulated against each ordinal position irrespective of the PAs.

A "Q" computed using Cochran's test for repeated observations (Hays, 1963) on these data was not significant ($Q = 1.35, \chi^2 > .05$). Also, each PA was separately taken and a frequency distribution was made against different serial positions in which it appeared. Of the six $\chi^2$ associated with the six PAs, none were significant ($p > .05$).

Since no differential effects of serial position on recall were obtained, no further analysis with the first item removed was undertaken (see Lovejoy and Farley, 1971).

**Discussion**

The results obtained in this study are in contradiction to earlier studies (i.e. Berlyne et al., 1966 and Farley and Lovejoy, 1968) which obtained significant facilitative effects due to stimulation with white noise. The recall for the no noise condition was not significantly different from that for the noise conditions. Further, there was no significant difference between the 75 dB and the 90 dB SPL conditions.
indicating neither facilitative nor detrimental effects of different levels of stimulation by white noise. However, considering the mean and mean percentages in Table I, it appears that the results may be confounded with the stimulus materials used in the PA task. The recall of the response terms was extremely poor in all cases suggesting that the learning task was extremely difficult. Runquist (1966) has pointed out that a difficult learning task may prevent subjects from attaining sufficient levels of learning to test hypotheses under consideration. Kleinsmith and Kaplan (1964) and Osborne and Farley (1971) used the same PAs and obtained very low levels of performance. Such floor effects make interpretation of the results rather difficult. Hence, the interpretations offered here are to be treated with caution.

Regarding the effects of timing of arousal, the results are again contradictory to both Farley and Lovejoy's (1968) study which obtained a reminiscence effect due to arousal stimulation between 3-6 minutes and the expectation from the electroconvulsive shock (ECS) and drug studies of decreasing recall as a function of time between learning and ECS or drug administration. In the present study, the results tend to indicate that stimulation during 3-6 minutes retards learning when compared to stimulation between 0-3 and 6-9 minutes. It may be noted that stimulation between 0-3 and 6-9 minutes cannot be interpreted as facilitative since the recall for the stimulation condition (noise condition) was not significantly better than the no stimulation (no noise) condition.

These results are difficult to interpret from the available literature. Considering notions proposed by Walker (1967), we could divide the course of a memory trace following the learning event into three functionally different time periods of short-term memory (STM), dynamic trace period
(DTP) and structural trace period (STP) or long-term memory.

From Walker's (1967) discussion, it appears that the putative STM duration is around 20-30 seconds followed by a longer period of action decrement phase (DTP). The STP appears to be characterized by the completion of the consolidation period with the trace attaining a relatively permanent strength, subject, however, to some forgetting with time. Although it may not be completely appropriate to relate the temporal parameters used in the present study to Walker's three trace periods, it appears that stimulation by white noise immediately following the learning and after the consolidation period is over, has little or no effect on the course or the strength of the memory trace, but any stimulation during the action decrement phase may be detrimental. It is possible that the DTP may have an initial pick-up phase which may last for one to two minutes following the STM phase. The results of the present study suggest that stimulation by white noise during the initial DTP pick-up phase may not have any effect. The second phase of DTP may be characterized by intense reverberation or consolidation activity during which period any stimulation may be harmful (3-6 minutes in the present study). The second phase of DTP is followed by the gradual fading of reverberation concomitant with the laying down of a relatively permanent trace (STP) during which periods arousal stimulation may have little or no effect on final trace strength.

The notions above are speculative, but are suggestive of further research seeking clarification of the temporal course of the human memory trace.
References


Footnote

Fig. 1: Experimental Conditions Representing Intensity and Temporal Parameters of Arousal
<table>
<thead>
<tr>
<th>Intensity of Arousal</th>
<th>Timing of Arousal</th>
<th>Combined Mean</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-3 min.</td>
<td>3-6 min.</td>
<td>6-9 min.</td>
</tr>
<tr>
<td>No White Noise</td>
<td>0.88</td>
<td>0.88</td>
<td>(14.58%)</td>
</tr>
<tr>
<td>75 dB</td>
<td>1.13</td>
<td>0.75</td>
<td>1.63</td>
</tr>
<tr>
<td>90 dB</td>
<td>1.13</td>
<td>0.50</td>
<td>0.88</td>
</tr>
<tr>
<td>Combined Mean</td>
<td>1.13*</td>
<td>0.63</td>
<td>1.25</td>
</tr>
</tbody>
</table>

*This includes only the scores for 75 dB and 90 dB groups and not the control group.
<table>
<thead>
<tr>
<th>Contrast Number</th>
<th>Intensity No Noise</th>
<th>75 dB 0-3</th>
<th>3-6</th>
<th>6-9</th>
<th>90 dB 0-3</th>
<th>3-6</th>
<th>6-9</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>&lt;1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>2.25</td>
<td>&gt;.05</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
<td>-2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>5.81</td>
<td>&lt;.025</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>&lt;1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>-2</td>
<td>1</td>
<td>-1</td>
<td>2</td>
<td>-1</td>
<td>&lt;1</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>1.97</td>
<td>&gt;.10</td>
</tr>
</tbody>
</table>