
EDRS Price MF-$0.65 PC-$3.29


In two separate paired-associate learning experiments each employing 60 university students as subjects, the contribution of individual differences (IDs) in arousal to short- and long-term retention was investigated using IDs in salivary response to lemon juice stimulation as an index of arousal. Experimental subjects were pre-selected out of 184 subjects on the basis of extreme arousal scores. The hypotheses were tested that high-arousal learning would lead to poor short-term retention but would demonstrate reminiscence or superior long-term retention relative to low-arousal learning; low-arousal learning was expected to lead to superior short-term retention but classical forgetting over the long-term relative to high-arousal learning. Experiment (Exp.) One yielded (non-significant) results in the predicted direction, whereas Exp. Two, incorporating procedural changes on the basis of Exp. One, confirmed the hypothesis (p<.025). (Author)
INDIVIDUAL DIFFERENCES
IN AROUSAL AND THEIR
RELATIONSHIP TO SHORT- AND
LONG-TERM RETENTION

CENTER FOR
COGNITIVE LEARNING
Technical Report No. 139

INDIVIDUAL DIFFERENCES IN AROUSAL AND THEIR
RELATIONSHIP TO SHORT- AND LONG-TERM RETENTION

John W. Osborne and Frank H. Farley

Report from the Project on Motivation and
Individual Differences in Learning and Retention
Frank H. Farley, Principal Investigator

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

March, 1971

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported
in part as a research and development center by funds from the United States Office of Education,
Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily
reflect the position or policy of the Office of Education and no official endorsement by the Office
of Education should be inferred.

Center No. C-03 / Contract OE 5-10-154
STATEMENT OF FOCUS

The Wisconsin Research and Development Center for Cognitive Learning focuses on contributing to a better understanding of cognitive learning by children and youth and to the improvement of related educational practices. The strategy for research and development is comprehensive. It includes basic research to generate new knowledge about the conditions and processes of learning and about the processes of instruction, and the subsequent development of research-based instructional materials, many of which are designed for use by teachers and others for use by students. These materials are tested and refined in school settings. Throughout these operations behavioral scientists, curriculum experts, academic scholars, and school people interact, insuring that the results of Center activities are based soundly on knowledge of subject matter and cognitive learning and that they are applied to the improvement of educational practice.

This Technical Report is from the Motivation and Individual Differences in Learning and Retention Project from Program 1. General objectives of the Program are to generate new knowledge about concept learning and cognitive skills, to synthesize existing knowledge, and to develop educational materials suggested by the prior activities. Contributing to these Program objectives, the Learning and Memory Project has the long-term goal of developing a theory of individual differences and motivation. The intermediate objective is to generate new knowledge of the learning and memory processes, particularly their developmental relationship to individual differences and to motivation.
# CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables and Figures</td>
<td>vii</td>
</tr>
<tr>
<td>Abstract</td>
<td>ix</td>
</tr>
<tr>
<td>I Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Arousal</td>
<td>1</td>
</tr>
<tr>
<td>Arousal and Verbal Learning</td>
<td>2</td>
</tr>
<tr>
<td>Arousal and Salivation</td>
<td>6</td>
</tr>
<tr>
<td>Methodological Considerations in Salivary Studies</td>
<td>7</td>
</tr>
<tr>
<td>The Purpose of this Study</td>
<td>8</td>
</tr>
<tr>
<td>II Experiment One</td>
<td>11</td>
</tr>
<tr>
<td>Method</td>
<td>11</td>
</tr>
<tr>
<td>Subjects</td>
<td>11</td>
</tr>
<tr>
<td>Materials and Equipment</td>
<td>11</td>
</tr>
<tr>
<td>Procedure</td>
<td>11</td>
</tr>
<tr>
<td>Results</td>
<td>12</td>
</tr>
<tr>
<td>Discussion</td>
<td>13</td>
</tr>
<tr>
<td>Summary</td>
<td>14</td>
</tr>
<tr>
<td>III Experiment Two</td>
<td>15</td>
</tr>
<tr>
<td>Method</td>
<td>15</td>
</tr>
<tr>
<td>Subjects</td>
<td>15</td>
</tr>
<tr>
<td>Materials and Equipment</td>
<td>15</td>
</tr>
<tr>
<td>Procedure</td>
<td>15</td>
</tr>
<tr>
<td>Results</td>
<td>15</td>
</tr>
<tr>
<td>Discussion</td>
<td>16</td>
</tr>
<tr>
<td>Summary</td>
<td>17</td>
</tr>
<tr>
<td>IV General Discussion</td>
<td>19</td>
</tr>
<tr>
<td>References</td>
<td>23</td>
</tr>
</tbody>
</table>
# LIST OF TABLES AND FIGURES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Total Recall Scores of Paired Associates as a Function of Arousal and Recall Interval</td>
</tr>
<tr>
<td>2.</td>
<td>Summary of Analysis of Variance of Recall Scores and Arousal Conditions</td>
</tr>
<tr>
<td>3.</td>
<td>Total Recall Scores of Paired Associates as a Function of Arousal and Recall Interval</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Distribution of Gross Salivation to Lemon Juice</td>
</tr>
<tr>
<td>2.</td>
<td>Differential Recall of High- and Low-Arousal Ss as a Function of Time</td>
</tr>
<tr>
<td>3.</td>
<td>Distribution of Net Salivation to Lemon Juice</td>
</tr>
<tr>
<td>4.</td>
<td>Differential Recall of High- and Low-Arousal Ss as a Function of Time</td>
</tr>
</tbody>
</table>
AROUSAL

Over the last 15 years the concept of arousal has undergone considerable change as new evidence has emerged from neurophysiological research. The attempt to define it in anatomical terms has evolved extensively because of the comparatively heavy balance of theory over actual knowledge in this area. Advances being made seem to indicate that early conceptions were oversimplifications of extremely complex phenomena. As Berlyne (1967) has indicated: "...at one time, arousal was commonly regarded as a clear-cut unitary variable, more or less identifiable with fluctuations in the degree of activity of the brain-stem reticular formation and with concomitant changes in electrocortical activity, circulatory activity, pupillary diameter, electrical and thermal properties of the skin, tension and activity of the skeletal musculature, and overall intensity of behavior" (p. 10). The fact that arousal is a common factor in a wide range of behaviors does not help us very much in specifying its role. Apparently almost any stimulus change except that to which a person is accustomed tends to cause changes in the above-mentioned indices of arousal. Additional evidence from anatomical data (Olszewski, 1954), drug studies (Anokhin, 1959), central stimulation (Fessard, 1960; Olds & Peretz, 1960) and ablation (Schiff, 1964, 1965) indicates that there is considerable differentiation of structure and function within the reticular formation. This increases the complexity of attempting to define the loci of arousal. Berlyne (1967) has reviewed some of the research which has accumulated to consolidate the growing impression that arousal is an extremely complex network of interrelated phenomena: the relevance of the interaction and balance between reticular and cortical activity (Hugelin & Bonvallet, 1957; Jouvet & Michel, 1958), the role of the hypothalamus (Gellhorn, 1961; Feldman & Waller, 1962), the limbic system (Maclean, 1949) and the specific or lemniscal sensory pathways (Sprague, Chambers, & Stellar, 1961) has been made apparent. There also appears to be evidence (Feldman & Waller, 1962; Wikler, 1952; Lacey, 1967) of at least three different kinds of arousal—"autonomic," "electrocortical," and "behavioral." Attempts to define arousal neurophysiologically have just begun to scratch the surface of a comparatively new area and as a result must be highly tentative depending on skillful theorizing to fill in the gaps left by current knowledge.

An alternative approach is to define arousal in molar fashion as does Duffy (1962). She defines the level of activation of the organism (arousal) as "...the extent of release of the potential energy stored in the tissues of the organism, as this is shown in activity or response [p. 17]." She maintains that the degree of activation of the organism is not synonymous with the degree of overt activity or the vitality of the organism or the availability of energy for response. It is rather, "...the extent of the release of stored energy of the organism through metabolic activity in the tissues [p. 18]." According to Duffy, a definition with more appeal for psychologists would be: "...activation is "...the arousal which occurs in the absence of physical exertion" (p. 19). At the present time the approach one takes to the concept of arousal is somewhat a matter of personal preference.

Because of the difficulty in defining arousal and the emerging fact that it involves a complex interrelationship of several variables, Berlyne (1967) has advised that it be regarded as a dimension rather than a phenomenon confined to one location in the central nervous system. Arriving at comparative measures of arousal is made difficult by individual differences (ID's) in the reactivity of the physiological processes involved in arousal (Lacey,
Bateman, & Van Lehn, 1953; Voronin & Sokolov, 1958; Schuore, 1959). The situation is further complicated by the fact that although changes in arousal are often simultaneously reflected in the so-called "indices of arousal," this is not always the case. Also, sometimes changes in these indices are anomalous—one increasing while the other decreases. The direction of the change depends upon the individual, the stimulating condition, and the phase of the experiment (Berlyne, 1960; Lacey, 1967). The conclusion to be drawn is that arousal level as a dimension needs to be related to the components of the situation in which it occurs before it becomes meaningful.

Writers such as Hebb (1955), Lindsley (1957), and Malmo (1958) have identified drive with arousal, making it the nuclear concept of motivation theory. However, as Berlyne has pointed out, the connection is not a simple isomorphic relationship. There is the problem of reconciling differentiation within drive to differentiation within arousal. There seems little doubt that few would object to the equating of general drive "...an increase in the overall activity level of the organism" (Berlyne, 1967, p. 17) to general arousal. The relationship of specific drives and the role in reinforcement to arousal is more open to question.

Hebb (1955) has argued that arousal plays an important role in reinforcement. He maintained that when arousal is below an intermediate optimal level, "...a response that produces increased stimulation and greater arousal will tend to be repeated" (p. 250) while excessive stimulation may be disruptive and repel. This notion was subsumed under Hebb's (1958) consolidation theory. Closely related to this view is the action decrement theory of Walker (1958) (described in detail elsewhere in this paper) which has also claimed a prime role for arousal in reinforcement. Berlyne (1960) has suggested that "...reward, even when it is associated with increases in stimulation, could depend on arousal reduction, so that a modified and extended drive reduction might be defended" (p. 24). Russia; writers such as Anokhin (1958), Polezhaev (1960) and Vinogradova (1959) have maintained that the orientation reaction (as manifested by the main signs of a transient rise in arousal) plays an essential part in learning. They have argued that it is crucial in the facilitation of the neural processes underlying learning. Maltzman and Raskin (1965) have recently advanced a similar view holding that "...elicitation of an orienting reflex constitutes a reinforcing state of affairs" (p. 15).

Berlyne (1967) has stated that there seems little to be gained from attempting to find specific physiological and neural centers which correspond to specific drives. The more useful and likely hypothesis seems to be that some common neurophysiological variable, such as general drive or arousal, underlies all forms of reinforcement.

AROUSAL AND VERBAL LEARNING

There appear to be similarities and differences in the way reinforcement operates in verbal and nonverbal learning. Contiguity alone, in verbal learning, clearly does not guarantee learning. Berlyne (1967) believes that the factors that reinforce verbal learning are those influencing "attention," a word that would seem to implicate arousal.

Over the last two decades there has been a steady growth in studies concerned with the relation between arousal and verbal learning. Research by Obrist (1950) and Thompson and Obrist (1964) has been concerned with relating the subjects' (SS) arousal level associated with individual items to the efficiency of learning a list. These authors measured galvanic skin response (GSR) and electroencephalographic (EEG) changes during serial learning of nonsense syllables. It was shown that while SSs were engaged in learning, mean GSR magnitude was higher than during control periods. The two indices (GSR and EEG) reflected a tendency for each syllable to produce the highest arousal at about the time it was beginning to be correctly anticipated. Obrist (1962) in another experiment with serial learning found correct anticipation on different days to be linearly related to heart-rate and electrodermographic measures of autonomic activity in two SSs and curvilinearly related in three SSs. Schönpflog (1963) has reported that during intentional learning as compared to incidental learning, skin resistance basal levels are lower and GSR incidence higher.

Schönpflog and Jeike (1964) found that exposure to a series of emotional words resulted in better immediate recall scores than exposure to a series of neutral words. This experiment also showed that the emotional words resulted in significantly higher basal conductance levels and GSR magnitudes.

Other experiments have studied the relationship between verbal learning and IDs in arousal. Berry (1962) recorded skin conductance during presentation of paired associates (PAs) under intentional learning conditions and found highest recall in SSs having intermediate conductance levels. Klein, Kaplan, and Tarte (1963) working with PAs obtained similar results when the interval between training and testing was 6 minutes but when the interval was increased
to 1 week, recall scores increased monotonically with skin conductance.

Some of the most dramatic experiments in the area of the relationship between arousal and verbal learning have been those carried out at the University of Michigan (Kleinsmith & Kaplan, 1963, 1964; Walker & Tarte, 1963). The rationale underlying this work has been based on the neural consolidation theory of Hebb (1958).

Support for this position has come from Kleinsmith and Kaplan (1963) in an experiment using a PA task with eight single words as stimulus items and eight single digits as response items. The stimulus words were KISS, RAPE, VOMIT, EXAM, DANCE, MONEY, LOVE and were judged on an a priori basis by the experimenters (Es) to be differentially arousing. A single learning trial was followed by a single recall trial. The interval between learning and recall was varied from 2 minutes to 1 week. In order to determine empirically the arousal effects of each stimulus word, skin resistance was recorded during learning. It was found that associations learned in the presence of low arousal, as indicated by little change in skin resistance, showed high immediate recall but poor delayed recall (the latter consisting of 45 minutes and 1 week) relative to items associated under high arousal as indicated by a large skin resistance change. The latter items demonstrated poor immediate recall but high recall on retention tests 45 minutes and 1 week later. This same type of result, i.e., poor immediate recall but higher later recall of items learned under high arousal, was later obtained by Kleinsmith and Kaplan (1964) using six 0% association nonsense syllables as stimuli and six single digits as responses.

Walker and Tarte (1963) replicate the Kleinsmith and Kaplan (1963) study using homogeneous and mixed lists of high and low arousal words. The high arousal stimulus words were MONEY, PAPE, SUIT, EMBRACE, KISS, VOMIT, PASSION, and SEX. The low arousal stimulus words were WHITE, POND, BERRY, FLOWER, WALK, PENCIL, GLASS, and CARROT. The response items were single digits. Three groups of Ss learned a low-arousal list; three groups learned a high-arousal list; and three groups learned a mixed list of half low- and half high-arousal words. Measures of skin resistance were taken during learning. Within each list, one group recalled the list at 2 minutes after learning, one group recalled at 45 minutes, and one group recalled at 1 week. Walker and Tarte found that the capacity to recall the number associated with low arousal words dropped as a function of time. The capacity to recall numbers associated with high arousal items dropped at 45 minutes and then rose slightly at 1 week. Although the magnitude of the effect was less than in the Kleinsmith and Kaplan studies, the results were statistically significant. Parrey (1968b) has recently used the stimulus words of the Walker and Tarte (1963) study in a free learning experiment and obtained results somewhat similar to those of Walker and Tarte (1963) and Kleinsmith and Kaplan (1963, 1964) with respect to the long-term recall measure. However, he did not obtain the cross-over effect between immediate and long-term recall.

The "action decrement" theory of Walker (1958) and the neural consolidation theory of Hebb (1958) used to explain the Michigan studies are adaptations of earlier versions of perseveration or consolidation theory. Although the theory emerged in the early 1900's, it fell into disrepute due to the lack of physiological evidence. However, as Glickman (1951) has pointed out, the accumulation of evidence now favors the existence of some mechanism of consolidation. Although alternative explanations can be advanced for many of the experiments supposedly supporting the existence of consolidation, similar conclusions have been drawn by Parrey (1968b) in a recent relevant review.

A commonly held hypothesis is that neural consolidation depends upon reverberating circuits. This idea stems from early work by Lorente de Nó (1938) and has been adopted in differing forms by Hebb (1949, 1958), Young (1953), and Gloor (1955).

The consolidation theory of Hebb used to explain the Kleinsmith and Kaplan studies implies that under conditions of low arousal, relatively little non-specific neural activity is available to support the reverberating trace, resulting in little consolidation and poor long-term retention. While under conditions of high arousal, the increased non-specific neural activity will result in more reverberation and consequently better retention. In the light of Hodgkin's (1958) finding that neurons are sharply limited in their maximum rates of firing repeatedly in a reverberating circuit, it is reasonable to expect that memory involving repeated firing of neurons will be at least temporarily inhibited. So it is that Walker (1959) and Parrey (1961) have hypothesized that during reactivation, while consolidation of the memory trace is occurring, the same trace is relatively unavailable to the organism, explaining why recall of material during the consolidation process is poor. Thus, under conditions of high arousal we have the slightly paradoxical situation of the consolidation process strengthening memory but...
until the termination of the process restricting its availability.

The "action decrement" theory of Walker (1958) as summarized by Walker and Tarte (1963) provides one of the best descriptions of the above phenomena:

(1) The occurrence of any psychological event, such as an effort to learn an item of a paired-associate list, sets up an active perseverative trace process which persists for a considerable period of time. (2) The perseverative process has two important dynamic characteristics: (a) permanent memory is laid down during this active phase in a gradual fashion; (b) during the active period, there is a degree of temporary inhibition of recall, i.e., action decrement (this negative bias against repetition serves to protect the consolidating trace against disruption). (3) High arousal during the associative process will result in a more intensely active trace process. The more intense activity will result in greater ultimate memory but greater temporary inhibition against recall [p. 113].

In commenting upon these predictions made by Kleinsmith and Kaplan on the basis of Hebb's neural consolidation theory, Eysenck (1967) has pointed out that they "...present a considerable extrapolation of existing knowledge" [p. 131] and that "...interest in consolidation has been minimal during the past thirty years, and consequently very little is in fact known about its precise modus operandi in these various fields of learning" [p. 113].

Another caution concerning the Michigan studies is that verbal responses will be reinforced most effectively when arousal is at an intermediate level. He points out that immediate recall will be governed by the interaction between the reinforcing effect of arousal and the effect of arousal on performance. He claims that these two effects can be expected to follow different inverted U-shaped functions. The interaction between learning and performance effects may provide the explanation of differences in long- and short-term recall. If the relationship between arousal and reinforcement is a U-shaped function, sometimes a monotonic increasing relation will appear while at other times a monotonic decreasing relation can be expected. The location of the individuals' position on the curvilinear function will depend on the interaction of both extra- and intra-individual determinants of arousal. In discussing results such as those of Kleinsmith and Kaplan it must be assumed that when recall was tested within 2 minutes of training, high-arousal items were further from the optimum level of arousal than low-arousal items, while for long-term recall, they were closer to the optimum.

So it can be shown that data, like those of Kleinsmith and Kaplan, following monotonic functions, can be compatible with hypotheses implying curvilinear functions, but the opposite is not true. As Berlyne (1967) has pointed out, "When recall scores have been obtained within a few minutes of training and compared for different subjects, several experimenters (Berry, 1962; Christ, 1962; Kleinsmith, Kaplan, & Tarte, 1953) have found them to be curvilinearly and nonmonotonically related to arousal measures, such as GSR reactivity and
heart rate. It is true that recall has sometimes increased with indices of arousal. Both average GSR magnitude and alpha frequency were correlated with learning rate in W. D. Obrist (1960) experiment. Increasing linear functions also connected skin resistance and heart rate in two out of five of P. A. Obrist's (1962) subjects. As already mentioned data following monotonic functions can be compatible with hypotheses implying curvilinear functions, but the opposite is not true. In Kleinsmith, Kaplan, and Tar. e's experiment, recall 1 week after training went up linearly with log-conductance level. It was therefore not inversely related to short-term recall as the Michigan theory would lead one to expect [pp. 69-70].

Eysenck (1967) has also shown interest in the inverted U-function between reinforcement and arousal. He has reported an experiment by Corcoran (1965) which attempted to relate the inverted-U relation to personality. Corcoran argued that introverts are higher on excitement arousal than extraverts, than the two groups should react differently to changes in experimental conditions. Corcoran used sleep deprivation as a deastring condition; incentives and loud noise as arousing conditions. Two tasks were used—one easy (multiple reaction time), and the other difficult (complex tracking). The results in each case showed differential behavior by extraverts and introverts. Corcoran concluded "(a) that the performance of the less-aroused subject deteriorates when the general level of arousal is decreased; (b) that the performance of the more aroused subject is less affected than that of the less-aroused subject when arousal generally is decreased and may even improve if the initial level of arousal is past the optimum" [p. 272]. In keeping with Berlyne's view, Eysenck points out that to increase or decrease arousal by stimulation depends greatly just where on the inverted-U function of arousal-performance a person is at the time of treatment. Like Berlyne, he sees the net result of arousal treatments as an interaction between the treatment and the intrinsic level of arousal within the subject. Eysenck (1967) has been led to believe that introverts are more arousable than extraverts. An individual's preferred "level of stimulation" will depend on his prevailing level of arousal so that introverts generally will prefer less arousal increase than extraverts. There is a considerable overlap between this line of thought and Berlyne's notions regarding "arousal potential."

There are a few experiments with designs that permit learning and performance effects to be separated. These studies have usually attempted to manipulate arousal in such a way as to preclude the confusing of the effects of general arousal level and the effects of the arousal elicited by particular words which are part of the learning task itself. Alper (1948) attempted to induce arousal by giving "ego-oriented" instructions (informing the Ss that the task was a measure of intelligence) on a PA list to one half of her Ss. The remaining half of the Ss were given standard "task-oriented" instructions. She tested for recall immediately after learning and 1 day later. "Ego-oriented" Ss not only recol led significantly more new items on Day Two than on Day One but also recalled on Day Two significantly more of the same items they had recalled on Day One than did the "task-oriented" Ss.

In a recent study employing drugs, Batten (1967) induced arousal by giving each of his Ss desoxanerin or phenobarbital prior to PA learning, by manipulating instructions to half the Ss so as to increase uncertainty and to promote "ego-involvement" (telling the Ss that the task was a measure of intelligence and that the Ss were going to find out how the Ss really operated) and by administering the Stroop Color test (Jensen & Rowher, 1966). The PA stimuli were words judged to be emotionally neutral: PAPER, AMONG, FAR, UPON, SUCH, MOST, BACK, and THAN. The responses were single digits. Following a single presentation of the lists, Batten tested for recall 2 minutes, 20 minutes, 45 minutes, 1 day, and 1 week later. Results were in the direction suggested by the Michigan experimenters but were not statistically significant.

King and his associates (Harper & King, 1967; King, 1963; King & Dodge, 1965; King & Walker, 1965; King & Wolf, 1965) have used a method of delayed auditory feedback to induce arousal. They have found that immediate retention of prose material practiced under delayed auditory feedback of 2 to .8 seconds is significantly poorer than that obtained from appropriate controls. However, on a long-term (24 hour) retention test, material practiced under delayed auditory feedback yielded greater retention, relative to the initial amount of material recalled, in comparison to the control group. In other words, the delayed auditory feedback group showed greater resistance to forgetting over the 24-hour period.

Berlyne, Borsa, Craw, Gelman, and Mandell (1955) and Berlyne, Borsa, Hamacher, and Koenig (1956) have induced arousal by using white auditory noise. The assumption that white noise is arousing is supported by the evidence that: white noise activates the recticular arousal system (Berrien, 1946; Costello & Hall, 1967; Gibson & Hall, 1969) and the finding that
continuous white noise causes skin resistance to drop significantly over a period of 15-20 minutes under conditions that would otherwise leave skin resistance virtually unchanged (Berlyne & Lewis, 1963).

Berlyne, et al. (1965), in the first of a series of PA experiments with white noise as the agent of arousal, used dysyllabic male first names as response terms and visual patterns as stimulus terms. They found that recall was impaired when Ss were administered 72 decibels (dbs.) of white noise during the two training trials and during the test trial 24 hours later. They suspected, however, that the results were in part due to the characteristics of the visual patterns used as stimuli.

In the second experiment, Berlyne, et al., (1965) used single dysyllabic adjectives (e.g., glassy) heterogeneous dysyllabic adjectives (e.g., glassy, crucial) as stimuli. The response terms were dysyllabic male first names. One-quarter of the items were learned under white noise and tested the next day under white noise; one-quarter, learned with white noise and tested without white noise; one-quarter, learned without white noise and tested under white noise, and one-quarter, learned and tested without white noise. Five groups of Ss received different intensities of white noise ranging from 35 dbs. to 75 dbs. They found that on the training day there was significantly less recall for items learned under white noise as compared to items learned with no white noise. On the test day 24 hours later, items learned under white noise the day before were recalled significantly more often than non-white noise items. No significant effect due to white noise during the test trial appeared. Variations in white noise intensity had no effect. On the basis of these two experiments, they concluded that white noise-induced arousal has a facilitative effect on learning rather than performance.

In a third PA experiment, Berlyne, Borsa, Hamacher, and Koenig (1966) again used single dysyllabic adjectives as stimulus terms and single dysyllabic male first names as response terms. Noise conditions were varied so that noise appeared only during the presentation of the stimulus, during the interval between items, during the presentation of the stimulus and response, or not at all. They found that white noise during presentation of stimulus and response terms in training trials significantly increased recall in a test trial given 24 hours later. Whether white noise was present or absent after the response made no significant difference on the 24-hour measure of retention. They also found that during training on Day One, white noise under all presentation conditions had no detrimental effect on recall. This finding is contrary to the previous findings of Berlyne, et al. (1965) and Kleinsmith and Kaplan (1963, 1964) in which arousal had a detrimental effect on immediate recall but enhanced long-term recall relative to the non-arousal condition. Farley and his associates (Farley 1968a & b, 1971; Farley & Gilbert, 1968; Farley & Lovejoy, 1969; Farley & Manske, 1969; Manske & Farley, 1970; Haveman & Farley, 1969; Jones & Farley, 1970, Gaa & Farley, 1969) have conducted an extensive series of studies on arousal and memory. Some of this work has been noted earlier, but it should be added here that they have successfully replicated the basic Kleinsmith and Kaplan (1963) and Walker and Tarte (1963) findings, although they concluded that these results seem to be more readily obtained with skin resistance and blood volume changes than with other systems and indices. In addition, they have found no effects of white noise on retention, except in free recall learning. Using an ID approach, however, they have found that IDs in arousal are measured either by salivation (see below) or personality are related to retention, although more strongly when salivation is used as an arousal index. The direction of the relationship is toward poorer short-term retention but superior long-term retention of high arousal Ss relative to low arousal Ss, in line with the Kleinsmith and Kaplan (1963, 1964) and Walker and Tarte (1963) results. In the first study to utilize an approach based on IDs in physiological arousal measured by salivation, Kindergartners were measured for salivary response during 20 secs. of stimulation by 4 drops of lemon juice, then stratified on the magnitude of response, and at a later date tested in a one-trial pictorial paired-associate task with either an immediate or 24-hour retention test (Farley & Gilbert, 1968). The interaction described above between arousal and retention was obtained. A notion of intrinsic arousal, that is, characteristic IDs in arousal, would seem to be a useful one where education is concerned, and will be returned to later.

AROUSAL AND SALIVATION

According to Sternbach (1966) each of the three salivary glands—the parotid, sublingual, and submaxillary glands—is linked by fibers to both the sympathetic nervous system (SNS) and parasympathetic nervous system (PNS). The effect of PNS stimulation is to increase the production of thin watery saliva. SNS effects are the opposite; salivary flow is decreased. Since this is true of each of the three glands, it follows that the total amount of saliva present in
the mouth represents the balance of SNS-PNS activity. More saliva indicates apparent PNS dominance; less saliva indicates apparent SNS dominance.

Sternbach (1966) has pointed out that arousal, activation, energy mobilization, and excitation/inhibition are synonyms used to describe the PNS/SNS balance. Eysenck (1953) has related the factorially derived personality dimensions of neuroticism and extraversion to autonomic balance. He has argued that neuroticism is reflected by deviation from autonomic balance either in the direction of SNS or PNS whilst extraversion and introversion are related to the direction of the deviation from autonomic balance. Those with apparent SNS dominance would be more extraverted while those with apparent PNS dominance would be more introverted. Hence, introverts should salivate more profusely than extraverts. Salivation should be an index of arousal or activation as well as personality.

The hypothesis has been advanced (Eysenck, 1963b, 1964, 1967) that introverts are characterized by a state of higher cortical arousal: "I have suggested a link between personality and the theory of excitation and inhibition by postulating that extraverts are characterized by the particularly rapid rise of cortical inhibition, its slow dissipation, and its relatively high level. Conversely, introverts are characterized by the slow growth, rapid dissipation, and generally low level of cortical inhibition. The opposite prediction is made with respect to excitation, that is, introverts show a high degree of excitation, extraverts a low degree" (1963b, p. 1033). It is important here to avoid the confusion of identifying behavioral inhibition with cortical inhibition. As Eysenck has stated, cortical inhibition, if anything, is more related to uninhibited than inhibited behavior. "By depressing the activity of the highest centres, cortical inhibition may give rise to uninhibited behavior" (Eysenck, 1963b, p. 1033).

There is some direct evidence to support this notion. Savage (1964) found a relationship between EEG alpha rhythm amplitude, inhibition, and extraversion. Extraversion, resulting from high cortical inhibition, resulted in significantly higher alpha rhythm than did introversion associated with low cortical inhibition. Marton and Urban (1966) also reported that habituation occurred more rapidly in extraverts than introverts. The evidence pointed to "inhibitory potential" developing faster in persons who tended towards extraversion even though average alpha frequencies at rest were lower for extraverts than for introverts. Farley (1968b) using the two-flash threshold and Shagas and Schwartz (1963) studying cortical evoked potentials have obtained similar results.

Eysenck (1967) has reviewed experimental evidence employing stimulant and depressant drugs indirectly supporting this notion. He has attempted to link psychological concepts such as "excitatory potential" and "inhibitory potential" with the physiological processes in the ascending reticular activating system. Drugs are known to change a person's position on the extraversion-introversion dimension in the direction of greater introversion (stimulant drugs) or in the direction of greater extraversion (depressant drugs). It has been shown that these drugs tend to act through the ascending reticular activating system (Eysenck, 1963a).

On the basis of Eysenck's hypothesis, it can be predicted that the effector output of a highly aroused organism would be greater than that of a lesser aroused organism, when both were subject to the same stimulation, if in fact, as Bremer (1954) believes, the neuro-physiological correlate of high levels of activation is a state of high cortical facilitation.

This deduction has been tested in several studies, including an experiment by Corcoran (1964) in which he showed that four drops of lemon juice placed on the tongue of introverts resulted in a significantly greater output of saliva than when the same stimulation was administered to extraverts. Eysenck and Eysenck (1967) were critical of this study because of the small number of Ss, the use of a relatively unknown personality inventory, and the unusually high correlations of personality variables with physiological measures. In a replication study, Eysenck and Eysenck (1967) confirmed Corcoran's findings by showing that introverted Ss secreted more saliva than extraverted Ss in response to the stimulus of four drops of lemon juice on the tongue. Because of this study and Corcoran's own replication, there seems to be good reason for serious consideration of the findings.

**METHODOLOGICAL CONSIDERATIONS IN SALIVARY STUDIES**

In a review of human salivary conditioning experiments, Feather (1965) has listed at least five different techniques for studying human salivary secretion. The absorbent technique, which was devised by Razran (1935), consists of placing a preweighed dental cotton roll under a S's tongue for a given time, reweighing and recording the difference increment between the two weights as the amount of saliva secreted. This technique has been
mouth movement is necessary if unwanted
motion of fluid to receptor cells. Movements will obviously affect the distribu-
tion of fluid to receptor cells. Controlling
mouth movement is necessary if unwanted

The establishment of representative individual
subjects is a prominent feature of studies using salivary response. The establish-
ment of representative individual and group rates of salivation is to some extent
arbitrary because it depends on the variables present at the time of measurement.
Whether one takes a particular part of the range of salivary responses such as mean or mode to be
representative will depend upon making certain a priori assumptions. The complex interrela-
tionships involved in physiological measurement make this more than just a statistical
problem.

The purpose of this study

In summarizing the literature concerning arousal and verbal learning, a wide range of
experiments employing arousal-producing stimulus terms, delayed auditory feedback,
drugs, frustrating tasks, and white noise suggest that arousal facilitates long-term recall.
The situation is obscure where the relationship between arousal and immediate recall is con-
cerned. Berlyne, et al. (1965), King and Dodge (1965), King and Wolf (1965), Klein-
smith and Kaplan (1963, 1964), and Walker and Tarte (1963) found arousal to have a sig-
ificantly detrimental effect on immediate recall. On the other hand, Alper (1948), Ber-
lyne, et al. (1966), Farley (1966a), Farley and Lovejoy (1969), and Haveman and Farley
(1969) (their free recall study only) found arousal to have no significant inhibiting ef-
fect on immediate recall but to increase long-
term recall relative to nonarousal conditions. Berlyne et al. (1966), and Haveman and Farley
(1969) in discussing this problem have sug-
gested that the effects of arousal may be de-
pendent on the nature of the learning material used.

Studies by Hörman and Todt (1960), Schönpläug and Schafer (1962), Schönpläug
and Belke (1964) (all cited in Berlyne, 1967), Stauffacher (1937), Courts (1939), and Malitz-
man, Kantor, and Langdon (1960) have shown increases in arousal during learning to produce
increases in immediate recall. On the other hand experimenters such as Berry (1962), P.A.
Obrist (1962), and Kleinsmith, Kaplan, and Tarte (1963) have found a curriculum and non-
monotonic relationship between immediate recall scores and arousal measures.
On the basis of available evidence, there appears to be some uncertainty regarding the relationship between recall and arousal. The object of the present research is to perform a modified replication of Kleinsmith and Kaplan's (1964) study but utilizing the research of Farley and Gilbert (1968) which studied intrinsic arousal rather than induced arousal, measuring IDs in the former by salivary response.
EXPERIMENT ONE

This experiment was concerned with the interaction of arousal and recall interval in nonsense syllable PA learning. It was similar to that of Kleinsmith and Kaplan (1964) except that arousal was treated as an individual difference (ID) variable, being measured in terms of salivary output to lemon juice (Eysenck & Eysenck, 1967) rather than being measured during learning trials in the form of, for instance, GSRs to each consonant-vowel-consonant (CVC) nonsense syllable, with perhaps some form of arousal manipulation (e.g., white noise).

The specific hypotheses to be tested were:

(i) Low arousal Ss would have significantly higher recall scores than high arousal S's over the short term (1.5 minutes).

(ii) High arousal Ss would have significantly higher recall scores than low arousal S's over the long term (24 hours).

METHOD

Subjects

The Ss were obtained from an undergraduate course in educational psychology at the University of Wisconsin; there were 46 females and 39 males. Participation in departmental research for up to 3 hours was a course requirement that was strongly urged, although not rigidly enforced.

Materials and Equipment

Stainless steel forceps, sterilizer, tongs, 1 c.c. glass syringe, lemon juice, standard cotton dental swabs, 50 test tubes with rubber stoppers, stop-watch, mirror, Right-a-Weight electronic balance, Kodak Carousel AV 900 projector, screen, Cousino Synco-Repeater Model SR-7341, 6 PA slides, 12 slides each with a different stimulus word only, and 28 slides of colored s's.

Procedure

A measure of salivary response to lemon juice was taken from each S by means of the absorbent technique (Razran, 1935). Standard cotton dental swabs were used throughout. Equipment coming into contact with the S's mouth was sterilized. The S was told that this measure was one of a series of physiological measures being taken in a study of individual differences. Each S was told that a standard dental swab would be placed under his tongue, with some harmless fluid being dropped onto the tongue which he was to hold there for 20 seconds. At the end of this interval he was told to simultaneously raise his tongue (for the swab to be removed) and swallow the fluid. A request was made for the mouth to be kept as widely open as possible during the operation for ease of access. The S was instructed not to make any attempt to manipulate the swab with his tongue. Before measurement began the experimenter demonstrated the two basic tongue movements involved; namely, touching the roof of the mouth with the tip of the tongue halfway back (for reception of the swab) and hollowing out the tongue (for reception of lemon juice). A mirror was provided to permit brief rehearsal by the S.

The swab was placed upon the sublingual salivary gland with forceps. Then four drops of lemon juice (0.1747 grams mean weight) were delivered to the tongue by means of a 1 c.c. glass syringe. In order to be sure of stimulating the "sour" taste receptors, the juice was dropped onto the lateral margins of the tongue allowing it to run towards the center.
At the end of 20 seconds the moistened swab was removed to be placed in a sealed test tube which had been previously weighed whilst containing the same swab in a dry state. The test tube and swab were weighed a second time, the difference between wet and dry weights constituting the amount of salivation to lemon juice. This operation was carried out with the utmost possible speed and precision. The S was seated so that the equipment tray was out of view. A cloth covered the tray when the S entered the room. Care was taken not to use the words "lemon" or "juice" or let the S have a close look at the syringe. Every effort was made to minimize distractions in the room in order to avoid spatial inhibition.

From the distribution of the salivary responses of 80 Ss (5 of the original Ss were lost for such reasons as being tongue-tied, inability to curl the tongue, etc.) the 20 highest scoring Ss (salivation > .40 grams) and 20 lowest scoring Ss (salivation < .10 grams) were selected for the second stage of the experiment.

The second, or learning, phase of the experiment was patterned after Kleinsmith and Kaplan (1964). The 20 high-arousal Ss and 20 low-arousal Ss were given a single learning trial with a list of six nonsense syllable-number pairs. (The following six 0% association value CVC nonsense syllables (Hilgard, 1951) were used: CEF, QAP, TO', JD(, LAJ, DAX. The response items were, respectively, the single digits from 2 to 7.)

A Kodak Carousel slide projector and Cousin timer were used to present the stimuli for an exposure time of 4 seconds. During the training trial the S first saw the nonsense syllable alone, and then repeated with a single digit response term. To separate the arousal effects of the stimuli from one pairing to the next, two slides containing four equidistant spots of four colors each were inserted before and after each PA and the S was instructed to name the colors (red, green, yellow, orange, black, and blue were used randomly on these slides). Two color-slides were presented prior to the first PA so that the S could "settle down" before the PAs were presented. The S was instructed to "concentrate carefully on both colors and nonsense syllable-number pairs," and to call them out loud, but to avoid rehearsal the S was not specifically told that he would be tested for recall.

During the recall session the S was instructed to recall the correct number for each nonsense syllable as it appeared for 4 seconds and to guess if uncertain. The correct numbers were not repeated. Colors were used as an interpolated task as before.

Ten high-arousal Ss and 10 low-arousal Ss were tested for immediate recall (1.5 minutes as measured from presentation of the first slide) while 10 high-arousal and 10 low-arousal Ss were tested for long-term recall (24 hours). The Ss within arousal levels were randomly assigned to retention intervals.

To correct for serial order effects, 10 different training lists were used, each list given to one S in each group. Two 6 x 6 balanced Latin squares were used to derive the lists after randomly omitting one row (Cochran & Cox, 1957). The order of the recall lists was varied in the same manner.

RESULTS

In Fig. 1 is presented the distribution of salivary responses to lemon juice of the 80 Ss. The cut-off points for the 20 highest and 20 lowest salivators were > .40 grams and < .10 grams, respectively.

Fig. 2 illustrates the differential recall of high- and low-arousal Ss as a function of time. At the immediate test, greater recall (percent correct) was demonstrated by the low arousal Ss than the high arousal Ss, whereas on the long-term test, the reverse was true. The trends represented in the figure are in the hypothesized directions. Table 1 summarizes the total recall scores for each of the four groups arising from arousal level—retention interval interactions.

In Table 2 is summarized the two way fixed effects analysis of variance Hays (1963) performed on the data to determine the significance of the trends evident in Fig. 2. It can be seen...
Fig. 2. Differential Recall of High- and Low-Arousal Ss as a Function of Time

Table 1
Total Recall Scores of Paired Associates as a Function of Arousal and Recall Interval

<table>
<thead>
<tr>
<th>Arousal Level</th>
<th>Recall Condition</th>
<th>Short-Term</th>
<th>Long-Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td></td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td>10</td>
<td>8</td>
</tr>
</tbody>
</table>

N = 10 for each group

Table 2
Summary of Analysis of Variance of Recall Scores and Arousal Conditions

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arousal Level</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Recall Condition</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Interaction</td>
<td>.4</td>
<td>1</td>
<td>.40</td>
<td>.68</td>
</tr>
<tr>
<td>Error</td>
<td>21.2</td>
<td>36</td>
<td>.58</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>21.6</td>
<td>39</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

that there was no significant effect due to arousal level recall condition or their interaction.

DISCUSSION

On the basis of Eysenck and Eysenck's (1967) finding that extraversion correlated equally well with salivation to lemon juice with or without the use of a base rate measure, basal salivation was not measured in this experiment. Salivary responses reported above referred to basal salivation plus net salivation due to lemon juice. The implication of the Eysencks' study is that basal salivation varies little from one individual to another or that the magnitude of its individual values is very low in comparison to stimulation levels of salivation. The distribution of salivary responses to lemon juice showed a pronounced positive skew. In practice this meant that it was more difficult to find Ss who secreted >.40 grams of saliva (high salivators).

Although an attempt was made to achieve precision and uniformity during salivary measurement, irregularities were inevitable. The operation of depositing and removing the swabs and depositing the lemon juice on the tongue were subject to inter-subject variation, e.g., inability to immediately hollow the tongue, narrow palates, failure to comprehend instructions, and variations in the time between deposition of the swab and the delivery of lemon juice to the tongue. In spite of these difficulties, few Ss were unable to hold the juice on their tongues for 20 seconds. The presence of a mirror helped greatly in this regard.

The hypotheses regarding the interaction of arousal state and recall were not confirmed; however, the data did show trends consistent with the hypotheses. As already mentioned, the procedure for selecting high- and low-arousal Ss may be a source of error responsible for the failure to obtain statistically significant results.

In both this study and that of Kleinsmith and Kaplan (1964), arousal has been generated from two sources: the stimulus materials and the intrinsic arousal level of the S. Kleinsmith and Kaplan measured the former in terms of GSR to each syllable, while ignoring the latter. In the present study the attempt has been made to discriminate between Ss on the basis of intrinsic arousal as measured by salivary output while ignoring stimulus-induced arousal. Both Eysenck (1967) and Berlyne (1967) have postulated a curvilinear relationship in the form of an inverted U between arousal and performance so that the effect of an arousal-inducing treatment will depend on the interaction between it and the S's intrinsic level of arousal. In both the above study and the present one, the arousal effects from these two sources may be confounded. The implicit assumption seems
to be that when examining IDs in one source of arousal, one can ignore the other source because it is randomly and normally distributed. However, it should be noted that the measurement of "intrinsic" arousal (stable IDs in arousal) represents itself a stimulating situation. That is, arousal is usually measured as a response to stimulation; thus it is obviously difficult to unambiguously differentiate between induced arousal and intrinsic arousal as discussed. The present study was an attempt to see whether the results of Kleinsmith and Kaplan (1964) could be reproduced when Ss were selected in terms of intrinsic, rather than induced, arousal. Because of the aforementioned difficulties relating to salivary measurement, results as spectacular as those of Kleinsmith and Kaplan were not expected; however, the procedural modifications suggested below should allow for a more precise test of the hypotheses. Possible the best solution to this problem is an experiment which selects Ss on the basis of intrinsic arousal and then specifically manipulates arousal. On the basis of the present study it appears that salivary response, as measured, is not as successful as PA-induced GSR in relating arousal to recall.

Perhaps this means that induced arousal will be reflected more in recall scores than intrinsic arousal.

During the training and recall trials, it was noticed that some Ss had difficulty in naming all four colors within 4 seconds. The equidistant arrangement of colored spots seemed to visually disorient them as they named colors in varying patterns from one slide to the next. Perhaps a straight line arrangement would alleviate this problem, allowing the eye its accustomed left-to-right lateral movement.

The fact that no S recalled more than two PAs suggests that the learning task was too difficult. After listening to instructions prior to recall, many Ss said that attempts at recall of the PAs would be entirely the result of guess. It is possible that a "floor effect" was present which prevented Ss from producing an adequate amount of learning to test the hypotheses under consideration (Runquist, 1966).

On the basis of Underwood and Schulz's (1960) two-stage theory of PA learning, it could be hypothesized that the Ss had insufficient time to accomplish both "response-learning" and "associative hook-up,"

Both Berlyne, et al. (1965), and Yerkes and Dodson (1908) contend that the effects of motivation upon learning are dependent upon the simplicity or complexity of the task. Suggestive support for this view comes from Haveman (1968) who found that the seemingly easier the verbal learning task the more sensitive the learning process became to differential effects of arousal. If the results of this study are interpreted in this light, there appears to be a case for making the task easier by increasing the exposure time for all slides during training and recall trials.

SUMMARY

From a total of 85 male and female Ss, 20 high-arousal and 20 low-arousal Ss were selected on the basis of salivary output to lemon juice stimulation. The Ss then participated in a PA-learning experiment in a 2 x 2 design with conditions of immediate and long-term recall, and high- and low-arousal, with 10 Ss per cell. It was hypothesized that: (i) high-arousal Ss would have significantly higher recall scores than low-arousal Ss over the long term; and (ii) low arousal Ss would have significantly higher recall scores than high-arousal Ss over the short term. Neither hypothesis was confirmed; however, the data demonstrated trends consistent with these hypotheses.
EXPERIMENT TWO

This experiment was a second attempt to study the interaction of arousal and recall interval in nonsense syllable PA learning. The hypotheses to be tested were the same as those of Experiment One. In light of the results obtained in Experiment One, a number of procedural modifications were made. These modifications were designed to reduce confusion due to the arrangement of colors in the interpolated task, increase the accuracy of the measurement of salivary response to lemon juice, reduce the possibility of floor effects in the recall scores due to task difficulty, and further reduce awareness of the main (retention) purpose of the experiment.

METHOD

Subjects

The Ss were 24 male and 75 female volunteers from an undergraduate learning course in educational psychology at the University of Wisconsin. Experimental participation requirements were identical to those of Experiment One.

Materials and Equipment

These were identical to those used in Experiment One except for the following: the colored spots on the "colored slides" were arranged equidistantly in a horizontal line. The choice and position of colors for each slide were random as before. Additionally, a Hunter GSR amplifier was used in the present study.

Procedure

The following changes were made to the procedure used in the first experiment: (i) Salivary response over 20 seconds to a dry swab was measured to establish basal salivation which was then subtracted from the salivation to lemon juice to give net salivary response to lemon juice. (ii) The exposure time for all slides was increased from 4 seconds to 6 seconds. (iii) Prior to learning and recall trials, Ss were attached to a disconnected GSR amplifier and told that a physiological measure would be taken while they were "doing a task." During learning and recall trials the E pretended to carefully watch the GSR amplifier which was placed behind the Ss. This procedure was used in an attempt to disguise the learning and retention nature of the experiment. (iv) The time between the beginning of the measurement of basal salivation and the beginning of the measurement of gross salivation to lemon juice was 2 minutes. (v) The time between the commencement of learning trials and immediate recall was approximately 2 1/2 minutes as a result of the increase in exposure time of all slides.

RESULTS

In Fig. 3 is presented the distribution of net salivation to lemon juice of 99 Ss. The cut-off points for the 20 highest and 20 lowest salivators were > .35 grams and < .025 grams, respectively. The increase in the accuracy of salivary measurement had the effect of truncating the range so the sample size was slightly enlarged in an attempt to offset this and make the extremes more apparent for selection as high- and low-arousal Ss.

The differential recall of high- and low-arousal Ss as a function of time is illustrated in Fig. 4. At the immediate test, greater recall (percent correct) was demonstrated by the low-arousal Ss than the high-arousal Ss, whereas on the long-term test, the reverse was true. The trends apparent in this figure are in the hypothesized directions.
In Table 4 is summarized the two-way fixed effects analysis of variance performed on the data to determine the significance of the trends evident in Fig. 4. There was no significant effect due to arousal level or recall condition, however their interaction was significant at the .025 level.

Table 4
Summary of Analysis of Variance of Recall Scores and Arousal Conditions

<table>
<thead>
<tr>
<th>Source:</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arousal Level</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Recall Condition</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>3.6</td>
<td>1</td>
<td>3.1</td>
<td>5.7*</td>
</tr>
<tr>
<td>Error</td>
<td>22.4</td>
<td>36</td>
<td>.624</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>26</td>
<td>39</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*_{2} < .025

DISCUSSION

The distribution of salivary responses to lemon juice had a marked positive skew. With most of the scores tightly clustered around a small part of the range it is apparent that IDs in salivation are going to be reflected in very slight differences in weight making the measurement procedure critical. The absorbent technique is subject to difficulties in this regard which when combined with the difficulty of temporal fluctuation of salivary secretion within SS makes reliable measurement of salivation a problem (Feather, 1965). The use of a ptygidal capsule device, a fixed head rest, mouth clamps to immobilize the mouth and control of SS's diet, activities, etc., prior to measurement are some of the increased controls which could be used in this regard. There is a need to try different time intervals between the measurement of basal and gross salivation as well as different amounts of lemon juice. Eysenck and Eysenck (1967) in their experiment do not state the time interval between measurement of basal and gross salivation or why the lemon juice was kept on the tongue for 20 seconds.

In spite of the procedural difficulties with the measures used, the significant interaction between arousal level and recall interval appeared here as it did in the work of Kleinsmith and Kaplan (1963, 1964); however, it was not as large as that found by these experimenters.
The present study also indicates the feasibility of an alternative measure of arousal in experiments concerned with arousal and recall and underscores the merit of ID analyses of the present type in learning and memory research. It would be of interest to combine salivary and GSR measurements in determining arousal level in future experiments. The investigation of the relationship between an intermediate level of arousal and recall was not pursued here because of a major preoccupation with verifying Kleinsmith and Kaplan's findings. However, this obviously needs future investigation.

As in Experiment One the range of scores for PAs correctly recalled was low. No S recalled more than two PAs. It seems desirable to design the learning task to permit a wider range of recall scores while at the same time maintaining control over associative characteristics or other features peculiar to the stimuli. It would be useful to take a random sample of Ss unselected with respect to arousal level and run them through the learning and recall procedure. This may help determine whether the low scores are more a function of the present S selection or the stimuli. A related possibility here would be inclusion of middle levels of measured arousal in future studies.

One problem which appears impossible to solve using the design of this experiment is that the stimuli when presented during recall will induce arousal which will affect the current arousal state of the S. One can only assume hopefully that this effect will be normally distributed over all subjects. It should be noted that Kaplan and Kaplan (1968), in reanalyzing the Kleinsmith and Kaplan (1964) recall data, found recall GSRs to be unrelated to retention.

As a result of this experiment and Experiment One, the evidence consistently indicates a relationship between arousal and recall like that found by Kleinsmith and Kaplan (1963, 1964). In spite of methodological problems salivation appears to be a feasible means of determining IDs in arousal level. Improvements in measuring salivation may lead to better opportunities for examining the arousal-recall relationship.

SUMMARY

From a total of 99 male and female Ss, 20 high- and 20 low-arousal Ss were selected on the basis of net salivary output to lemon juice stimulation. The Ss then participated in a PA-learning experiment in a 2 × 2 design with conditions of immediate and long term recall, and high- and low-arousal, with 10 Ss per cell. It was hypothesized that: (i) high-arousal Ss would have significantly higher recall scores than low-arousal Ss over the long term; and (ii) low-arousal Ss would have significantly higher recall scores than high-arousal Ss over the short term. Both hypotheses were confirmed at the .025 level of significance.
IV
GENERAL DISCUSSION

The results of this study demonstrate that the relationship between level of arousal and recall, as reported by Kleinsmith and Kaplan (1963, 1964), can be obtained in a similar paradigm when arousal is determined by salivary response to lemon juice. However, the interpretation of this relationship poses a problem: It may be argued that at least two kinds of arousal were involved both in this study and that of Kleinsmith and Kaplan. These were induced arousal resulting from exposure of the S to stimulus words and the intrinsic or chronic arousal level of the S in a learning test situation. The above authors, although not explicit on this point, apparently assumed that differences in intrinsic arousal level were randomly and normally distributed over the Ss used in their experiment and consequently not responsible for any effects on recall. Arousal effects were assumed to be entirely the result of experimental induction. Kleinsmith and Kaplan may have assumed that because arousal effects seemed to be fairly evenly distributed over the six stimuli for the three time intervals used, the same state of affairs would hold true regarding intrinsic arousal of the Ss. The present study has shown that the distribution of salivary responses to lemon juice is strongly skewed in a positive direction. If this measure is a valid index of intrinsic arousal, then it seems probable that a sample of undergraduates is going to contain a preponderance of Ss with a low level of intrinsic arousal.

The question of the interaction between these two sources of arousal bears greatly on the interpretation of Kleinsmith and Kaplan’s findings. One might argue that the two remain confounded in their experiment. Unanswered questions remain, such as: what proportion of the high- or low-arousal GSRs to the stimulus words were due to the arousal-inducing properties of the stimulus words or the intrinsic arousal of the Ss? The stimuli were simply classified as high- or low-arousal words on the basis of the Ss’ responses to them, making their arousal properties entirely relative. To say that, nevertheless, the Ss were in a high- or low-arousal state in terms of their gross reactions to the stimuli is true, but leaves unexplained the dynamics of the interaction between induced and intrinsic arousal.

The current study appears to have shown that there are individual differences in intrinsic arousal which are differentially related to recall. The question of the arousal effects of the stimulus material was answered by Kleinsmith and Kaplan (1964) who found “no systematic trends...in the distribution of items which could account for differences in behavior of high- and low-arousal learning” (p. 125). On this basis in the present study it was assumed that differences in recall were due to intrinsic arousal alone. Again, however, one cannot rule out interactions between the two sources of arousal in the determination of recall. Furthermore, the argument that salivary response may reflect induced arousal, with the lemon juice and testing situation representing arousal-inducing stimulation, rather than intrinsic arousal also cannot be entirely ruled out. The problem of obtaining unambiguous measures of intrinsic arousal may be a formidable one.

The value of the present research where arousal theory and human memory are concerned depends on the demonstrated reliability and validity of salivation as an arousal measure. Ferley and Osborne (1969) have obtained stability estimates over 24 hours of .81 for basal salivation, .78 for gross salivation to lemon juice, and .78 for net salivation to lemon juice (all at p < .01). Although no data are available on the very long-term stability of individual differences in salivation as measured here, it is interesting that the estimates reported by Ferley and Osborne are as high as those usually found in such areas of differential
psychology as aptitude and intellectual measurement. As regards the validity of salivation as a measure of arousal, Farley and Osborne reported a correlation of -.57 (p < .02) between salivation and the threshold of fusion of paired light flashes, or two-flash threshold, which has been shown through correlations with skin conductance (Maley, 1967) and EEG alpha amplitude (Venables & Warwick-Evans, 1967) to be a measure of arousal. The relationship between salivation, as measured, and two-flash threshold is not an exceptionally strong one in absolute magnitude, however, where arousal is concerned it is an encouraging one in an area not noted for significant relationships among measures (Sternbach, 1966). Taken in conjunction with the results of the two experiments reported in this paper, it reinforces the construct validity of salivation as a measure of individual differences in arousal.

As a result of changes made in the visual arrangement of the interpolated stimuli, as well as changes in the time interval and measurement of salivation, the differences between total group scores for high- and low-arousal Ss under the various periods of recall increased in Experiment Two. It is not possible to say whether one or all of these changes were responsible. The actual range of scores remained low pointing to continued difficulty with the stimulus material. There appears to be no way available at present to make recall of the stimulus material easier without introducing factors which will be unique to the stimulus and less directly comparable to the Michigan work.

A basic criticism of the design used in this study and those of Kleinsmith and Kaplan (1963, 1964) has been that arousal-inducing stimuli (nonsense syllables) were presented on both learning and recall trials. This meant that at the time of recall the Ss were exposed to the same words whose initial arousal-inducing function was the object of investigation, thus confounding the arousal effects of the two presentations. This objection, if valid, would have made consolidation between learning and recall impossible to evaluate and disallowed a specification of the temporal point of action of arousal in the memory sequence. In answer to this criticism, Kaplan and Kaplan (1968) re-analyzed the data of Kleinsmith and Kaplan (1964) to find that GSRs at recall did not correlate with recall performance. They concluded that recall GSRs did not predict recall performance while the learning GSRs did.

A unique feature of the present research has been the attempt to measure arousal as an ID variable. As mentioned earlier, it is impossible to completely avoid the possible arousal-inducing stimulation involved in the measurement of arousal by means of salivation. If salivation is a valid approach to determining arousal level one might hypothesize it puts us nearer the source of arousal itself by enabling direct measurement of reaction to a standard stimulus rather than reaction to stimuli, such as nonsense syllables, which sometimes increase and sometimes decrease arousal level because of the confounding of intrinsic and induced arousal. The stimulus of lemon juice always increased salivary output so that arousal could be measured in a unidirectional way for simple quantification.

This involved one measurement, whereas the method of Kleinsmith and Kaplan consisted of taking GSR deflections to six stimulus words which were ranked and then dichotomized to determine high and low levels of arousal.

Research appears to indicate that arousal can play an important part in the learning and memory process. Just what specific effect arousal has upon learning under varying conditions is still not known. Therefore, premature widespread application of learning theories derived from research on arousal should be avoided at this stage. The transition to the classroom needs to be made with care if we are to avoid costly blunders. Hilgard's six steps (Hilgard, 1966) involved in the gradual progression from research to widespread educational practice might be an appropriate model to follow. There is still need for a great deal of further programmatic research in this area as well as some exploratory studies which attempt to assess the effect of arousal in applied learning situations.

The most significant finding in the arousal literature, for education, is that arousal can facilitate long-term retention. The results of the current study are further evidence of this.

If the arousal levels of individuals can be manipulated to facilitate long-term retention, there are obvious economies which can be made in time spent on learning. Much depends, however, on the success of the means of inducing arousal. The method most often used (white noise) has had variable results and involves the repeated use of costly equipment for each individual.

What might be considered as a pilot study on the application of the findings of the Michigan studies to an educational setting was carried out by Levonian (1967). The essential change in experimental conditions was that the arousal-inducing material to be recalled was presented continuously (a driver education film) rather than as discrete stimuli as separated by fixed time intervals. It was
found that material presented during high arousal (as indexed by GSR) showed poor short-term retention and enhanced long-term retention. Under the condition of low arousal the reverse appeared to be true. Unfortunately, Levonian employed the same Ss for both short- and long-term measures. There is an obvious need for further experiments in this area to assess the feasibility of applying the results of laboratory research to applied educational settings as well as determining whether the effects replicate outside the laboratory.

Although measuring salivation is unlikely to affect arousal level significantly and hence not facilitate long-term retention it could provide a basis for establishing individual differences in arousal which could be utilized when programming instructional sequences. For example, the effect of an arousal-inducing agent may depend on the intrinsic arousal level of the S. Consequently, knowledge of an individual's level of intrinsic arousal may be an ID variable or "entering behavior" (Glaser & Reynolds, 1966) which is as useful as a knowledge of his intelligence when instructional techniques or regimes are being planned.
REFERENCES

Alper, T. G. Task orientation and ego orientation as factors in reminiscence. *J. exp. Psychol.*, 1948, 35, 224-238.


Gellhorn, E. Prolegomena to a theory of emotions. Perspectives in biology and medicine, 1951, 4, 403-436.


Maltzman, I., & Raskin, D. C. Effects of individual differences in the orienting reflex on conditioning and complex processes. *J. exp. res. in pers., 1965, 1*, 1-16.


Staufacher, J. C. The effect of induced muscular tension upon various phases of the learning process. J. exp. Psychol., 1957, 21, 26-46.
Vinogradova, O. S. [The role of the orientation reflex in the process of establishing conditioned connections in man.] In E. N. Sokolov (Ed.) Orientirovchny refleks i voprosy vyshei nervnoi i zatelnosti v norme i patologii. [The orientation reflex and questions of higher nervous activity in normality and pathology.] Moscow: Acad. Pedag. Sci., 1959.
Voronin, K. A., & Sokolov, E. N. [Electroencephalographic study of the conditioned connection with extinction of the orienting reflex in man.] Zh. vys. nerv. Deiat., 1958, 8, 3-16.