JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR 4, 291-299 (1965)

Effects of Stimulus Complexity and Induced Arousal on Paired-Associate Learning[1]

D. E. BERLYNE, DONNA M. BORSA, MARGARET A. CRAW, ROCHEL S. GELMAN, AND ELIZABETH E. MANDELL

University of Toronto, Toronto, Canada

Three experiments on paired-associate learning were carried out, with disyllabic male first names as response terms. Exps. I and II, in which visual patterns served as stimulus terms, showed significantly fewer correct recalls when stimulus terms were more complex or incongruous. Familiarization of half of the stimulus terms before the first training trial made no appreciable difference, which casts doubt on the attribution of the effect to differences in arousal value. The effect seemed more likely to be due to presence or absence of symmetry.

White noise was used in Exp. II as a means of manipulating arousal. Recall was impaired when *Ss* were subjected to 72 db white noise during training and testing, and there was some indication that recall might be improved by 58 db white noise.

In Exp. III, stimulus terms were single adjectives, homogeneous double pairs of adjectives, and heterogeneous double pairs. One quarter of the items were learned under white noise and tested the next day under white noise, one quarter learned without white noise and tested under white noise, one quarter learned under white noise and tested without white noise, and one quarter learned and tested without white noise. Five groups of *Ss* had different intensities of white noise ranging between 35–75 db. It was found that recall, both on the training day and on the test day, was worse with heterogeneous double than with homogeneous double stimulus terms but equal with homogeneous double and single stimulus terms, corroborating the hypothesis that the symmetry factor had been responsible for the results of Exps. I and II. On the training day, there was significantly less recall with white noise than without white noise. On the test day, items learned under white noise the day before were recalled significantly more often than others, although there was a significant interaction between presence of white noise on training and test days. No significant effect of white noise during the test trial appeared. Variations in white-noise intensity had no significant influence.

Theoretical discussions of learning have hitherto given a great deal of consideration to the biological and motivational significance of unconditioned or reinforcing stimuli and to the discriminability of conditioned or cue stimuli. There are, however, accumulating indications (Berlyne, 1960, 1963a, 1964) that motivational properties of conditioned or cue stimuli, and in particular properties that influence attentiveness and arousal, may be of decisive importance also.

This article reports three experiments concerned with effects on paired-associate verbal learning of level of arousal and of stimulus properties that are usually covered by words like "complexity," "incongruity," and "irregularity." Previous experiments have shown these properties (represented by the patterns displayed in Fig. 1) to influence duration

[1] This investigation was supported by grant APT-75 from the National Research Council of Canada and by research grant MH-06324 from the National Institute of Mental Health, U.S. Public Health Service. The data of Experiment I were reported by Miss Craw in a thesis submitted in accordance with the requirements for the degree of M. A. of the University of Toronto.

PP1 (11).max

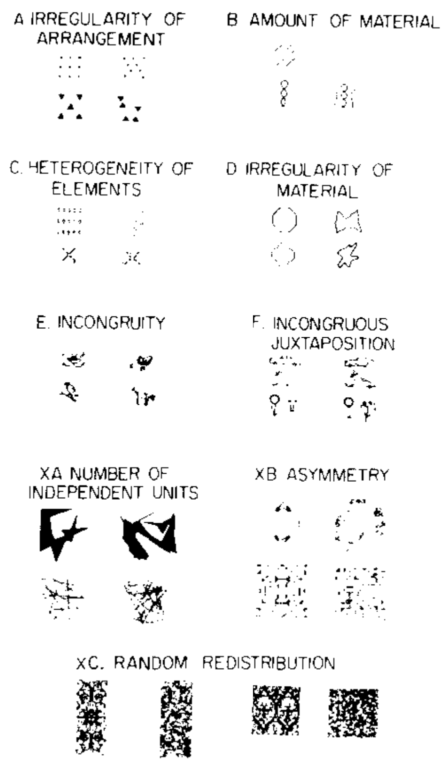

FIG. 1. Patterns used as stimulus terms in Exps. I and II.

of exploration (Berlyne, 1957, 1958; Berlyne and Lawrence, 1964), exploratory choice (Berlyne, 1963b), "interestingness" and "pleasingness" ratings (Berlyne, 1963b), amplitude and incidence of the GSR (Berlyne, Craw, Salapatek, and Lewis, 1963), and the duration of EEG desynchronization (Berlyne and McDonnell, 1965).

In two of the experiments to be reported, white noise was used as a means of manipulating arousal level. An earlier experiment (Berlyne and Lewis, 1963) had shown that white noise raises one index of arousal, namely skin conductance, and keeps it raised for at least 10–15 min.

EXPERIMENT I

Aims

The patterns in Fig. 1 consist of 18 pairs, each comprising a "less complex (LC)" member, displayed on the left, and a "more complex (MC)" member, displayed on the right. The pairs are, however, divided into nine categories, and the actual variable distinguishing LC and MC patterns differs from one category to another. Since these variables appear, in the light of previous experiments, to influence several aspects of behavior in generally similar ways, their effects will be analyzed collectively as well as separately.

While the experiment was being planned, we learned of the results obtained at the University of Michigan (Kleinsmith and Kaplan, 1963, 1964; Walker and Tarte, 1963), according to which responses coupled with stimuli productive of high arousal (judged by GSR amplitude) are recalled worse immediately after learning than those coupled with stimuli productive of low arousal but better when intervals ranging from 45 min to one week have elapsed since learning. Since our MC patterns had tended in a previous experiment (Berlyne, Craw, Salapatek, and Lewis, 1963) to evoke more frequent or more intense GSRs than our LC patterns, it was decided to bring *Ss* back for an additional recall test after 24 hrs to see whether a similar interaction would appear.

Subjects

There were 20 male and 20 female *Ss* from elementary psychology classes meeting during the academic year. Half of the *Ss* in either sex group had already undergone another paired associate learning experiment, which will not be reported here and which involved associating letters of the alphabet with pictures of animals, on the day before and on the same day.

Material

The patterns used as stimulus terms are shown in Fig. 1. The response items were familiar disyllabic male first names (e.g., ROBERT, WILLIAM).

Two sets of stimulus-response pairs were assembled and used with equal numbers of *Ss*. The stimulus terms of set A consisted of the LC member of one pair and the MC member of the other pair in each of the nine categories of Fig. 1. These patterns were arranged in four random orders, and each was randomly coupled with a response term. The remaining patterns constituted set B. Four sequences of set B were constructed by replacing each stimulus term in set A with the other member of the same pair in Fig. 1. Corresponding LC and MC stimulus patterns thus occupied the same temporal position and were coupled with the same response term. The different halves of the subject sample.

Procedure

The *S* went through four training trials, each using a different one of the four randomized sequences, on the first day. On each of these trials, a stimulus term appeared, projected on a screen, first alone for 4 sec, and then for a further 3 sec with the response term printed underneath. There was then an interval of 7 sec before the next stimulus term appeared. The second, third, and fourth of these trials gave *S* an opportunity to anticipate the response terms before they actually appeared, and so these are designated test trials 1, 2, and 3.

About 24 hrs. later, *S* underwent a fourth test trial, in which only stimulus terms were projected, each appearing for 4 sec, with an interval of 4 sec between it and the next. The sequence was the same for all *Ss* having set A, and the corresponding sequence, with counterparts from Fig. 1 substituted, was used for all *Ss* taking set B.

The *S* was told that this was an experiment on learning, designed to find out how people in general learn rather than to study individuals. When a name appeared, he was to pronounce it. On the second and third training trials, he was to try to recall and pronounce the name while the stimulus item was visible alone, but, as soon as the name appeared, he was to pronounce it whether or not he had correctly anticipated it. Before test trial 4, he was simply instructed to try to recall and pronounce the name that had been coupled with each stimulus term as it appeared.

Results

Overall LC/MC Comparisons. The mean numbers of response terms associated with LC and MC stimulus items that were recalled correctly are shown in Table 1. There were more recalls with LC items on every trial, although the difference was very slight on test trials 3 and 4. Analysis of variance shows the LC-MC difference to be significant for

TABLE 1

MEAN NUMBERS OF ITEMS RECALLED IN EXP. I

	Day 1			Day 2	
	Test trial 1	Test trial 2	Test trial 3	Test trial 4	Total
LC items	4.7	7.3	8.3	8.3	28.6
MC items	4.0	6.8	8.2	8.3	27.3
Total	8.7	14.1	16.5	16.6	55.9

the three test trials on the first day, $F(1, 114) = 10.3$, $p < .005$. Scores varied significantly from trial to trial, $F(2, 76) = 327.5$, $p < .001$. The interaction between trials and the LC/MC variable was, however, not significant. On test trial 4, held one day later, the LC/MC difference was not significant.

Category-by-Category LC/MC Comparisons. The LC mean exceeded the MC mean in eight out of nine categories (all except E), which is significant at the .05 level by the sign test. Wilcoxon T-tests reveal that the LC-MC difference reaches significance only in categories A (irregularity of arrangement, $T = 77$) and XC (random redistribution, $T = 76$). In both cases, $p < .02$.

Discussion

It looked, therefore, as if the expectations derived from the findings of the Michigan group had been half verified. More responses to LC stimulus terms (believed to be low-arousal stimuli) were recalled. This difference was most clearly in evidence in the first two trials that embodied a test for recall. There was, however, no significant interaction between the LC/MC variable and trials, and there was no sign whatever of a cross-over making for better recall with MC items after a delay of 24 hrs.

EXPERIMENT II

Aims

One way to confirm that the LC-MC difference obtained in Exp. I was really due to a difference in arousal value between LC and MC stimulus patterns is to give *Ss* a chance to inspect some of the patterns in both classes before learning begins. So Exp. II had a familiarization phase making use of the "Button-pressing Task" that had figured in previous experiments on exploratory behavior (Berlyne and Lawrence, 1964; Berlyne and Lewis, 1963): *S* is simply given control of the button operating an automatic projector and allowed to look at each of the stimulus patterns in turn for as long as he wishes. According to hypotheses propounded elsewhere (see Berlyne, 1960, 1963a), the initial appearance of a particular pattern will induce perceptual curiosity, conceived as a form of heightened arousal

PP1 (11).max

or drive, and S will continue to inspect the pattern until prolonged exposure to it has reduced curiosity to a threshold value. Even without accepting these hypotheses, one may conclude, on the basis of a copious literature (see Berlyne, 1960), that a stimulus pattern that has been experienced in the recent past will be less arousing that one that is novel.

If the LC-MC difference is due to differences in arousal value, one would expect (1) better recall of responses to stimulus terms that have already been seen in the familiarization phase, and (2) an attenuation of the LC-MC difference with stimulus terms that have already been seen.

A further objective of Exp. II, pursuing the influence of arousal on this kind of learning, was to sample the effects on recall of white noise.

Subjects

There were 96 *Ss*, comprising 4 males and 92 females, all of whom were students attending elementary psychology courses during the summer session.

Twenty-four *Ss* heard white noise at an intensity of 72 db (measured with a sound-pressure meter) through earphones, and 24 heard 58 db white noise. The remaining 48 received no white noise, although they wore the earphones for the purpose of control, being told that these would "help to exclude extraneous sounds."

Procedure

The procedure was exactly the same as that used in Exp. I with the following exceptions: (1) The whole procedure was completed in one session, during which S underwent three training trials (the second and third of which constituted test trials 1 and 2) and, immediately afterwards, a trial (test trial 3) for which only stimulus terms were presented (as on test trial 4 in Exp. I). (2) Before training began, there was a familiarization phase during which S saw nine of the 18 stimulus terms projected in turn on the screen. The button controlling the projector was placed in his hand. He was told that the patterns he was about to see were some of those that would figure in later parts of the experiment. He was to look at each pattern for as long as he wished and then to press the button, whereupon the pattern would be replaced by another.

Sets A and B of stimulus terms were used with equal numbers of *Ss* in the two white-noise groups and the no-white-noise group. For the familiarization phase, the stimulus patterns to be used with each subgroup were divided into two sub-sets, shown to equal numbers of *Ss* within the subgroup. One sub-set consisted of five LC and four MC patterns, all from different categories and arranged in a random order, except that LC and MC items were distributed as evenly as possible between the two halves of the sequence. The other sub-set consisted of the remaining four LC and five MC items, each occupying the same temporal position as the pattern belonging to the same category in the first sub-set.

Results

Overall LC/MC Comparisons. Table 2 shows the mean numbers of correct recalls with LC and MC stimulus terms. It will be noted that, as in Exp. I, there are more correct anticipations with LC items. Of the 96 *Ss*, 58 made more correct anticipations with LC than with MC items, 31 made more with MC than with LC, and there were seven ties, so that $p < .01$ according to the sign test.

There was nothing approaching a significant difference between pairs whose stimulus terms had been seen during the familiarization phase and those whose stimulus terms had not been seen. Nor was there any indication of a significant interaction between this variable and the LC/MC variable.

Category-by-Category LC/MC Comparisons. The LC-MC difference was significant, according to Wilcoxon T-tests, in categories C (heterogeneity of elements, $T = 603$, $p < .01$), XB (asymmetry, $T = 903$, $p < .05$) and XC (random distribution, $T = 744$, $p < .05$) only.

TABLE 2

MEAN NUMBERS OF ITEMS RECALLED IN EXP. II

	Test trials			
	1	2	3	Total
LC items	3.6	5.6	7.1	16.3
MC items	3.2	5.2	7.0	15.4
Total	6.8	10.8	14.1	31.7

Effects of White Noise. Significant interactions between white-noise intensity and other independent variables being absent, the recall scores for the 72 db, 58 db and no-white-noise groups are displayed in Table 3. It will be seen that the 58-db group scored better, and the 72-db group worse, than the no-white-noise group. The scores for the three groups differ significantly at the .01 level, $F(2,90) = 6.93$. Tukey's multiple-comparison test (Tukey, 1949) reveals that the 72-db group differs from the no-white-noise group at the .01 significance level and from the 58-db group at the .001 level, but that the latter two groups do not differ significantly from each other.

TABLE 3

MEAN NUMBERS OF ITEMS RECALLED AS A FUNCTION OF NOISE LEVEL (EXP. II)

	Test trial			
	1	2	3	Total
No white noise (48 *Ss*)	7.0	11.2	14.6	32.8
58 db white noise (24 *Ss*)	8.2	12.8	15.8	36.8
72 db white noise (24 *Ss*)	4.9	8.4	11.3	24.6

Discussion

The predictions that there would be better recall with familiarized stimulus terms and that there would be a smaller LC-MC difference with familiarized stimulus terms were not confirmed. This casts grave doubt on the assumption that the LC-MC differences obtained in Exps. I and II were due to differences in arousal value between LC and MC stimulus terms. The phenomenon seems, in fact, not to be closely related to the one demonstrated by the Michigan group.

It is, admittedly, conceivable that the subject-controlled exposure times (generally of the order of 5 sec—see Berlyne and Lawrence, 1964) in the familiarization phase were not sufficiently long to reduce arousal appreciably, but data in the literature strongly suggest that marked habituation occurs after a few seconds of exposure to a stimulus pattern. There might even be differences between LC and MC patterns in the readiness with which they evoke mediating or labeling responses. However, a more promising hypothesis is indicated by an examination of those categories in which significant LC-MC differences emerged in the category-by-category analyses of either or both experiments. These are categories A, C, XB, and XC. In all of these categories, the LC patterns, in contrast with their MC counterparts, contain a number of identical features (or features that are identical except for mirror-image inversions). The hypothesis that this factor was responsible for the finding thus suggests itself.

As regards the effects of white noise, we may conclude that it impairs recall at 72 db but that this detrimental influence is absent at 58 db. We cannot conclude that 58 db white noise improves recall, in view of the failure to obtain a significant difference between the 58-db and no-white-noise groups.

EXPERIMENT III

Aims

The first aim of Exp. III was to test the hypothesis suggested by the findings of the last two experiments, namely that recall is worse when compound stimulus terms contain unlike elements side by side than when they contain identical elements side by side. Since a satisfactory test requires application of the hypothesis to a different kind of stimulus material, a memory drum was used in place of the automatic slide projector and, while the response terms still consisted of disyllabic male first names, disyllabic printed adjectives replaced the visual patterns as stimulus terms. Some items had ***heterogenous double*** stimulus terms (e.g., "glassy crucial") and others had ***homogeneous double*** stimulus terms (e.g., "crucial crucial"). In case the predicted superiority of recall with homogenous double, as compared with heterogenous double, stimulus terms were verified, it would be important to know whether recall is improved by the former or depressed by the latter. For this reason, items with ***single*** stimulus terms (e.g., "crucial") were also used, so that a baseline could be established.

A second aim, indicated by the contrasting effects on recall of 58 db and 72 db white noise in Exp. II, was to sample a wider range of white-noise intensities. Five intensities, ranging from 35 db to 75 db in steps of 10 db, were accordingly used with different groups of *Ss*.

Finally, it was desirable to separate the effects on performance. By an "effect on learning," we mean a difference in probability of recall between items learned under different conditions but tested under identical conditions. By an "effect on performance," we mean a difference in probability of recall between items learned under identical conditions but tested under different conditions. To distinguish such effects, a design is needed in which the presence or absence of white noise during training and the presence or absence of white noise during testing are factorially separated. This is the kind of design that has become standard in the animal literature, in connection, for example, with the effects on learning and performance of conditioned-stimulus intensity (cf. Kessen, 1953), drive level (cf. Lewis and Cotton, 1957), and light increment following a motor response (cf. Berlyne, Salapatek, Gelman, and Zener, 1964).

Subjects

There were 90 *Ss*, all of them female undergraduates from elementary psychology classes meeting during the academic year.

Design

Four half-lists of nine items, consisting of three items with heterogeneous double, three with homogeneous double, and three with single stimulus terms, were constructed. Care was taken to avoid adjectives that would clearly be recognized as applicable to human beings. These half-lists comprised Version I of half-lists 1, 2, 3 and 4, respectively. Version II was formed out of Version I in the following manner: all heterogeneous double items were turned into homogeneous double items by omitting the first adjective and repeating the second one, homogeneous double items were turned into single items by presenting the adjective once, and single items were turned into heterogeneous double items by placing a second adjective in front of the one already present. The same transformations were used to make Version III out of Version II. One third of the *Ss* in each of the ten subgroups had each of the three versions.

Half of the *Ss* (Group A) first learned a list comprising half-list 1 followed by half-list 2 while subjected to white noise (see Table 4). They went through three training trials with this list, the items of each half-list being in a different randomized order on every trial. The same procedure was then followed with a list consisting of half-lists 3 and 4 while the white noise was absent. The next day, there was a test trial, on which the items in each half-list appeared in a fourth randomized order. This time, half-list 4 came first without white noise, followed by half-lists 3 and 2 with white noise, and then by half-list 1 without white noise. Group B went through a sequence of events (see Table 4) with the necessary counterbalancing to control for order of presentation of white noise and no white noise and of the different half-lists.

TABLE 4
DESIGN OF EXP. III

Group A: Three training trials (Day 1)	Half-lists	Test trial (Day 2)	Half-lists
White noise	1	No white noise	4
	2	White noise	3
No white noise	3		2
	4	No white noise	1
Group B: Three training trials (Day 1)	**Half-lists**	**Test trial (Day 2)**	**Half-lists**
No white noise	2	White noise	3
	1	No white noise	4
White noise	4		1
	3	White noise	2

Within each group, subgroups of nine *Ss* had white noise at 35, 45, 55, 65, and 75 db, respectively.

Procedure

During the three training trials given on the first day, the stimulus term first appeared alone towards the left of the memory-drum aperture for 4 sec and then reappeared with the response term to its right for 4 sec. Blank paper then appeared for 8 sec, whereupon the stimulus term of the next item was presented. During the test trial given on the second day, each stimulus term appeared alone for 4 sec and there was a 4-sec interval before the appearance of the next stimulus term. Otherwise, the same procedure was followed as in Exps. I and II. The *Ss* were warned before white noise was due to begin and when it would be discontinued.

Results

Effects of Stimulus Terms. During the second and third training trials, held on Day 1, the mean numbers of correctly recalled items were 4.0 for heterogeneous double stimulus terms, 4.5 for homogeneous double stimulus terms, and 4.5 for single stimulus terms, $F(2,850) = 4.7$, $p < .01$. The corresponding means for the test trial held on Day 2 were 4.2, 5.0, and 5.0: $F(2,935) = 7.0$, $p < .001$.

Effects of White Noise on Recall during Training Trials (Day 1). Table 5 shows the mean numbers of correct anticipations during the second and third training trials held on Day 1. Recall was significantly worse for items learned under white noise, $F(1,850) = 6.8$, $p < .01$. The mean for trial 3 was significantly higher than the mean for trial 2, $F(1,85) = 135.4$, $p < .001$. The interaction between trials and presence or absence of white noise was insignificant. The effect of white noise-intensity and all the interactions in which this variable figured were insignificant.

Effects of White Noise on Recall during Test Trial (Day 2). Table 5 shows mean recall scores during the test trial held on Day 2. Recall was significantly better with items that had been learned under white noise on the previous day, $F(1,935) = 5.6$, $p < .025$. The difference made by the presence or absence of white noise during the test trial was not significant, but the interaction between this variable and the presence or absence of white noise during the previous day's training trials was significant at the .005 level, $F(1,935) = 10.2$. Effects and interactions involving white-noise intensity were all insignificant.

Discussion

The prediction that heterogeneous double stimulus terms would produce worse recall than homogeneous double stimulus terms received clear-cut confirmation on both the training day and the test day. The suspicion that the symmetry factor was responsible for the LC-MC difference in Exps. I and II is therefore corroborated. The finding that single stimulus terms produced mean scores that were very close to those produced by homogeneous double stimulus terms indicates that recall was depressed for heterogeneous double items rather than enhanced for homogeneous double items. The results of Exp. II suggest that this is not likely to be due to a difference in arousal value. Presumably, it has something to do with simultaneous processing of distinct items of information coming in through parallel channels in the nervous system or with distribution of attention, but more than that can hardly be said at the present stage.

TABLE 5
MEAN NUMBERS OF ITEMS RECALLED IN EXP. III

Condition during training trials	Training trials (Day 1): Training trial 2	Training trial 3	Total	Test trial (Day 2), Condition during test trial: White noise	No white noise	Total
White noise	4.9	7.5	12.4	3.5	4.0	7.5
No white noise	5.3	8.4	13.7	3.7	3.0	6.7
Total	10.2	15.9	26.1	7.2	7.0	14.2

The data for the test day provide an answer to the question of whether white noise affects learning or performance: recall on the test day was significantly improved when items had been learned under white noise the day before, but the presence or absence of white noise on the test day made little difference. If arousal is to be equated with drive, this finding is rather interesting in view of the contrast with what has been found in animal experiments. These have regularly shown marked effects of drive on performance, but indications of an effect of drive on learning have been scanty and slight (Kimble, 1961).

There was, however, a significant interaction, such that items learned under white noise but tested without white noise or vice versa, i.e., items that had undergone a change

PP1 (11).max

from learning to test conditions, had better recall than unchanged items. In the current state of knowledge, it would require some temerity to offer an explanation of this interaction. Every item was a changed and an unchanged item for equal numbers of *Ss*. On the other hand, changed items constituted the first and third half-lists (i.e., they appeared earlier on the average than unchanged items) on the training day and the second and fourth half-lists (i.e., they appeared later on the average than unchanged items) on the test day. Sequential effects could thus conceivably have played a part.

The finding that white noise during learning impaired recall on training trials but improved recall on the test trial held a day later is consonant with the findings reported by the Michigan group (Kleinsmith and Kaplan, 1963, 1964; Walker and Tarte, 1963), according to which high-arousal stimulus terms give rise to worse recall than low-arousal stimulus terms immediately after learning but to better recall after longer delays.

Our results represent, however, an advance in two respects. First, we manipulated arousal through white noise, whereas the Michigan group simply distinguished high-arousal and low-arousal stimuli according to the magnitude of the GSRs they evoked. This leaves open the possibility that GSR magnitude was correlated with some other factor that was actually responsible for the differences in recall. Secondly, in the experiments by the Michigan group, the stimuli raising arousal were the stimuli with which the responses to be learned were to be associated. In our experiment, arousal was raised by a completely independent agent, belonging, in fact, to a different sensory modality from the material figuring in the learning process. For these two reasons, we can have greater confidence that we are dealing with effects of arousal as such and not with some factor (e.g., some aspect of associative content) correlated with GSR magnitude or with an effect peculiar to arousal-raising properties of a cue stimulus. At the same time, the remarkable similarity between our findings and those of the Michigan group increases our confidence that white noise affects recall through arousal.

It will be remembered that, in Exp. II, 72 db white noise significantly impaired recall during and immediately after the training trials, while 58 db white noise improved it to a non-significant degree. In Exp. III, however, variations in white-noise intensity had no significant effects. There was, however, some hint of curvilinearity, with maximum effects at about 65 db in both cases. It seems plausible that recall may actually be nonmonotonically related to intensity of arousal with the form of the function dependent on the nature of the material to be processed.

References

BERLYNE, D. E. Conflict and information-theory variables as determinants of human perceptual curiosity. *J. exp. Psychol.*, 1957, **53**, 399-404.

BERLYNE, D. E. The influence of complexity and novelty in visual figures on orienting responses. *J. exp. Psychol.*, 1958, **55**, 289-296.

BERLYNE, D. E. *Conflict, arousal and curiosity.* New York: McGraw-Hill, 1960.

BERLYNE, D. E. Motivational problems raised by exploratory and epistemic behavior. In S. Koch (Ed.), *Psychology—a Study of a Science*, Vol. 5. New York: McGraw-Hill, 1963 (a).

BERLYNE, D. E. Complexity and incongruity variables as determinants of exploratory choice and evaluative ratings. *Canad. J. Psychol.*, 1963, **17**, 274-290 (b).

BERLYNE, D. E. Emotional aspects of learning. *Ann. Rev. Psychol.*, 1964, **15**, 115-142.

BERLYNE, D. E., CRAW, M. A., SALAPATEK, P. H., AND LEWIS, J. L. Novelty, complexity, incongruity, extrinsic motivation and the GSR. *J. exp Psychol.*, 1963, **66**, 560-567.

BERLYNE, D. E., AND LAWRENCE, G. H. Effects of complexity and incongruity variables on exploratory behavior and verbally expressed preference. *J. gen. Psychol.*, 1964, **71**, 21-45.

BERLYNE, D. E., AND LEWIS, J. L. Effects of heightened arousal on human exploratory behavior. *Canad. J. Psychol.*, 1963, **17**, 398-410.

BERLYNE, D. E., AND McDONNELL, P. Stimulus complexity and duration of EEG desynchronization. *Electroenceph. clin. Neurophysiol.*, 1965, **18**, 156-161.

BERLYNE, D. E., SALAPATEK, P. H., GELMAN, R. S., AND ZENER, S. L. Is light increment really rewarding to the rat? *J. comp. physiol. Psychol.*, 1964, **58**, 148-151.

KESSEN, W. Response strength and conditioned stimulus intensity. *J. exp. Psychol.*, 1953, **45**, 82-86.

KIMBLE, G. A. *Hilgard and Marquis' Conditioning and Learning*, 2nd Ed. New York: Appleton-Century-Crofts, 1961.

KLEINSMITH, L. J., AND KAPLAN, S. Paired-associate learning as a function of arousal and interpolated interval. *J. exp. Psychol.*, 1963, **65**, 190-193.

KLEINSMITH, L. J., AND KAPLAN, S. Interaction of arousal and recall in nonsense syllable paired-associate learning. *J. exp. Psychol.*, 1964, **67**, 124-126.

LEWIS, D. J., AND COTTON, J. W. Learning and performance as a function of drive strength during acquisition and extinction. *J. comp. physiol. Psychol.*, 1957, **50**, 189-194.

TUKEY, J. W. Comparing individual means in the analysis of variance. *Biometrics*, 1949, **5**, 99-114.

WALKER, E. L., AND TARTE, R. D. Memory storage as a function of arousal and time with homogeneous and heterogeneous lists. *J. verb. Learn. verb. Behav.*, 1963, **2**, 113-119.

(Received June 9, 1964)