Journal of Educational Psychology
1973, Vol. 65, No. 3, 304-312

FREE RECALL OF CATEGORIZED AND UNCATEGORIZED LISTS: A TEST OF THE JENSEN HYPOTHESIS[1]

ARTHUR R. JENSEN[2] AND JANET FREDERIKSEN
University of California, Berkeley

Jensen's two-level theory of mental abilities (Level I: rote learning and memory, Level II: abstraction and conceptual learning) was examined in terms of Level I and Level II learning tasks (free recall of uncategorized and categorized lists, respectively) administered to white and Negro second and fourth graders. Performance measures were amount recalled and amount of clustering in recall of the categorized list. The results accorded with previous studies based on other Level I–Level II tests given to low- and middle-socioeconomic-status children (both white and Negro) in different age groups, namely, a larger socioeconomic-status (or Negro–white) difference on Level II than Level I measures and the difference increasing with age.

Free recall of categorized and uncategorized lists can be used to test Jensen's hypothesis of two fundamental types of mental abilities and their interaction with social class and race. Jensen's hypothesis posits two broad classes of mental abilities called Level I (rote learning and memory) and Level II (intelligence, that is, analytical understanding, reasoning, abstraction, and conceptual thinking). Level I tests essentially call for accurately registering sensory experiences, immediately giving already well-learned names or labels to these, and at some later point in time repeating these labels in response to some partial stimulus cue. Level II tests, on the other hand, involve transformation and mental manipulation of the input in order to produce the answer. This may consist of relating and comparing present stimuli with past learning, generalizing and transferring old learning to a novel problem, or abstracting conceptual and semantic similarities and differences. This formulation and the evidence on which it is based have been presented in detail elsewhere (Jensen, 1968, 1969, 1970a, 1971b, 1973).

A number of previous studies have found a distinctive relationship between these two classes of ability and socioeconomic status, both within white groups and across racial (white and Negro) groups. The most consistent finding is that low- and middle-socioeconomic-status groups (and representative Negro and white groups) differ less, on average, on measures of Level I than of Level II ability (Guinagh, 1971; Jensen, 1968, 1969, 1970a, 1971b, 1973; Keogh & Macmillan, 1971; Orn & Das, 1972; Rapier, 1968; Wallace, 1970).

Another finding, though somewhat less consistently reported, is that the correlation between Levels I and II is lower in low-socioeconomic-status than in middle-socioeconomic-status groups; one consequence of this is that low-IQ (i.e., Level II) children of low-socioeconomic status perform better on Level I tests than their low-IQ middle-socioeconomic-status counterparts (Jensen, 1970b).

Based largely on White's (1965) research on the development and hierarchical arrangement of associative and cognitive abilities, roughly corresponding to Jensen's Levels I and II, it has been hypothesized that Level I and Level II abilities have dif-

[1] This research was supported by the U.S. Office of Economic Opportunity Contract OEO 2404.

[2] Requests for reprints should be sent to Arthur R. Jensen, Institute of Human Learning, University of California, Berkeley, California 94720.

ferent growth curves from infancy to maturity, with Level II coming into prominence later in development. If Level II processes account for less of the ability variance in younger than in older children, one should expect a relatively smaller socioeconomic-status difference in Level II in younger than in older children. So far there has been little systematic investigation of this hypothesis concerning the developmental aspects of Levels I and II.

The method of free recall can be used to investigate some of these aspects of the Level I–Level II hypothesis. Jensen (1971b) has claimed that

> Associative clustering in verbal free recall is one of the clearest forms of evidence of conceptual, hierarchical processes. For clustering to occur, the subject must actively organize the stimulus input according to certain self-provided superordinate categories [p. 61].

In other words, the processes involved in associative clustering in free recall closely correspond to Jensen's characterization of Level II ability. Lists can be made to differ in the degree to which they elicit clustering, that is, the tendency for subjects consistently to recall items in clusters of several conceptually or semantically related items. A categorized list elicits the most clustering. Such a list is composed of a number of items all of which can be classified into a smaller number of more general categories, such as furniture, food, animals, etc. As shown originally by Bousfield (1953), when college students are given a categorized list, with the items presented sequentially in the completely random order, the items are later recalled in clusters corresponding to the superordinate categories. The amount of recall is related to the tendency to cluster, and many studies by Bousfield and others have established that free recall of categorized lists is superior to the recall of uncategorized lists, at least for young adults. Uncategorized lists are composed of items that are associated with one another only remotely, if at all, and cannot be grouped in any obvious superordinate categories. Thus the acquisition and recall of an uncategorized list is more likely to involve rote learning and possibly some organization involving simple pairwise associations based on primary stimulus generalization, clang association, or functional relationship (e.g., ball–apple, box–fox, shoe–foot). Therefore, learning an uncategorized list would tend more to involve Level I ability. Note, however, that the learning of a categorized list need not depend upon Level II ability; it can be learned in the same Level I fashion as an uncategorized list. In the case of the categorized list, the learner's approach, whether predominantly Level I or Level II, depends upon which ability he brings to bear on the task. This will be largely a function of the subject's age and his rate of development in Level II ability.

The relevance of free recall to various aspects of the Jensen hypothesis has been investigated in three doctoral dissertations. In all three, representative samples of Negro and white children were compared; the groups differed in socioeconomic status, and there was no attempt to match the groups in this respect, since the studies were concerned with differences in learning characteristics between typical pupils from white and Negro neighborhoods, as in the present study. Some of the findings from the studies were (*a*) no white–Negro difference in kindergartners in amount of free recall or amount of clustering, (*b*) a large difference between kindergartners and fifth graders (both races), (*c*) a large difference between Negro and white fifth graders in free recall and amount of clustering (in favor of the white group), (*d*) substantial correlations between mental age (Peabody Picture Vocabulary Test) and amount of clustering in both groups at fifth grade but not at kindergarten (Glasman, 1968), (*e*) similar comparisons in the fifth and seventh grades showing a smaller white–Negro difference in seventh than in fifth grade for amount recalled, with the clustering score significantly favoring the white group in both grades (Gerdes, 1971), (*f*) a difference in recall and clustering for third and sixth grades, with the white–Negro difference greater for clustering than for amount recalled, and (*g*) the special training in categorization and class inclusion concepts raising the overall performance of all groups, but a significant white–Negro difference in clustering in the sixth grade (Peterson, 1972).

The present study carries this line of research further by comparing low- and middle-socioeconomic-status second and fourth graders on uncategorized and categorized lists, the latter presented under random and block conditions. The blocked presentation (serially grouping the items by category) should help to cue and faciliate clustering and recall.

The main experimental outcomes expected from the Jensen hypothesis are best described in terms of a nested design, with the racial (or socioeconomic-status) differences nested within the younger and older age groups and within recall tasks (uncategorized vs. categorized vs. blocked). In the younger age group (second graders in the present study) there should be relatively little white–Negro difference in performance on either the Level I test (recall of the uncategorized list) or on the Level II test (recall and clustering measures on the categorized list), since the two racial groups are hypothesized to be similar in Level I ability and in younger children the racial groups are not yet very differentiated in Level II ability. In older children (fourth graders in the present study) the racial groups should show little or no difference in recall of the Level I uncategorized lists, but should show a divergence in the Level II measures, that is, relatively better recall and more clustering of the categorized list by the white group. In other words, there should be little racial difference in the Level I measures in either age group, while on the Level II measures the racial difference should be magnified in the older age group. There is no theoretical expectation for the blocked condition. It is included to see if clustering in free recall can be primed or inculcated by this form of presentation in children who are unlikely to cluster spontaneously.

There is no reason to believe that having to learn categorized or uncategorized lists should differentially affect the subjects' motivation during the testing situation. It has been claimed, for example, that the standard intelligence tests may arouse anxiety in some children, thereby depressing their performance, or that some children may simply "turn off" and not try on a test that looks too difficult or unfamiliar to them. But categorized and uncategorized lists, when the items are matched for word frequency, look much alike to subjects, especially if the same group of subjects is not getting both kinds of lists. The instructions and procedures for both are identical, so it seems most unlikely that categorized and uncategorized lists should elicit any different test-taking attitudes or motivational states.

Method

Subjects

Representative samples of age-matched Negro and white second and fourth graders, 120 in all, were selected from two public schools, one in a predominantly Negro and the other in a white neighborhood. The neighborhoods involved and the parents' occupations listed in the school registers indicate that the white and Negro groups differ in average socioeconomic status, the white group being middle class and the Negro being predominantly lower-middle and lower class in terms of the usual criteria for judging socioeconomic status. But does this not confound race and socioeconomic status?

Our aim was to obtain quite typical samples of the two racial groups as they are found in this locality, without matching for socioeconomic status, since the interest was in studying Level I—Level II differences between samples that are quite typical of population groups whose average difference in scholastic performance has been the subject of national concern. Matching on background variables that are correlated with the dependent variables of the study but that are not known to be *causally* related to the dependent variables could only obscure the nature of the cognitive differences between the groups. If one equated groups on *enough* of the factors that happen to be correlated with the dependent variable, then presumably the groups would be practically equated on the dependent variable as well, and the null hypothesis would never be rejected. The interest is not on minimizing the existing differences by partialing out correlated (but not necessarily causal) variables, which prejudges the nature of the differences, but in pinpointing the nature of the children's performances in these groups just as they are, in the present case, in terms of Level I and Level II processes. Adequate description must precede the search for causal explanation; when causes are not already established or experimentally controlled, matching or partialing can only obfuscate description. For this reason the practice has been justifiably termed the "partialing fallacy" or "sociologist's fallacy."

At the descriptive stage of investigation, when experimental control is not feasible, explicit confounding of variables is highly preferable to creating the false illusion that the causal contributions of the variables are unconfounded by matching on one of the variables or partialing it out statistically.

Ten white and 10 Negro children in each grade were randomly assigned to one of three conditions: (*a*) an uncategorized list, (*b*) a random, categorized list, and (*c*) a blocked-categorized list. It should be noted that no subject was tested in more than one of the experimental conditions.

Procedure

Subjects were tested individually by the second author. Each subject was presented with a set of 20 familiar objects and was told he would have to remember and recall the names of all the objects he was shown. The objects were presented sequentially at approximately a 2-second rate, and subjects named the items as they were presented. (All the items were easily familiar and named by all subjects.) Each object was always removed from view before the next was presented. When all 20 items had been shown and named by the subject, he was given 90 seconds to recite the names of all the objects he could recall. This procedure was repeated for five trials. Instructions to the subject and all other features of the testing procedure were exactly the same for the three kinds of lists. The lists were matched for Thorndike-Lorge word frequency.

The *uncategorized* list consisted for the following toy objects, presented in a different random order on each trial: ball, bell, book, box, brush, car, chair, clock, coat, cup, egg, flag, frog, gun, horse, key, pen, spool of thread, train, and wheel.

The *categorized* list consisted of items representing four categories: clothing, tableware, furniture, and animals. The objects, presented in a different random order on each trial, were coat, dress, hat, shoe, skirt; cup, glass, plate, spoon, knife; bed, chair, dresser, lamp, table; and mouse, chicken, dog, horse, cow.

The *blocked-categorized* list consisted of the same items as the categorized list, but all the items of one category were always presented in sequence. The items were presented in a different random order within category blocks on each trial, and the order of the category blocks was varied randomly on every trial.

The experimenter recorded the subjects' responses and their order of emission on a specially prepared form.

Results and Discussion

Since not all readers may be acquainted with the kind of analysis applied to the present results, a word of explanation is in order. Ordinarily one might first think of the obvious factorial analysis of variance model for the analysis of this experiment, that is, a $2 \times 2 \times 3$ factorial design (Race $\times$ Grade $\times$ List). One would then test the three main effects and all their interactions. Statisticians have assured us, however, this factorial analysis of variance model is inappropriate for the variables and hypotheses under consideration in this study. What is clearly called for is a nested or simple effects design. Since the statistical rationale for the choice of this model in cases such as the present study has been thoroughly explicated by Marascuilo and Levin (1970), it is not necessary here to explain it in detail. The point is that the hypotheses under consideration do not involve interactions in the sense in which these are defined (and consequently tested) in the factorial analysis of variance model, which is best suited to experiments involving two (or more) *experimental* (i.e., manipulated) variables, A and B, and their interaction, $A \times B$, with random assignment of subjects to all cells. Such an $A \times B$ interaction is defined as a component that involves *every* cell of the design. When "interactions" are hypothesized, which are not really interactions in the sense of the analysis of variance model, but are merely differential effects within rows or columns, the nested design is called for instead of the factorial design. The present hypotheses are of this nature, that is, they involve differential effects between the white and Negro groups in learning some tasks (Level II) but not others (Level I) and at some age levels (older) but not others (younger), where tasks and age levels are the 3×2 conditions in which the racial differences are nested. The nested design is much better suited to localizing the conditions (i.e., the particular conjunction of rows and columns) in which the hypothesized effect occurs. Unlike the factorial model, it tests specific differences within conditions and does not test for interactions as in the factorial analysis of variance, which assumes that the interaction variance is equally distributed among all cells contributing to the interaction. The nested design permits the overall testing of one or more of the main effects in the usual way,

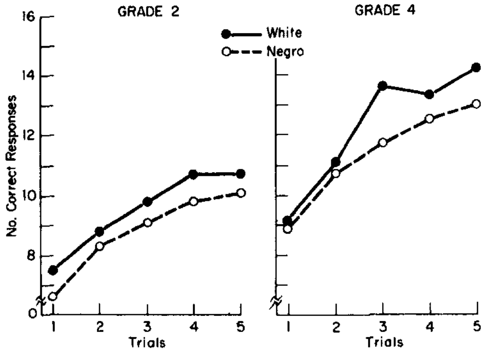

FIG. 1. Amount of free recall of random, uncategorized 20-item list.

but then, instead of testing interactions, deploys the sums of squares and degrees of freedom in such a way as to test the simple effects within each set of conditions. In some conditions no differential effects are hypothesized. These simply "water down" the analysis of variance interaction and can often completely obscure the hypothesized effect. The nested design, on the other hand, tests whether there are significant effects precisely where they are hypothesized to occur. In the present analyses, race is nested in grades and treatments, and grades are nested in treatments. (Other forms of nesting are, of course, possible, but they may not be as directly relevant to the hypotheses of interest.)

Since we are dealing with a repeated-measures design (i.e., five learning and recall trials), in which there is little interest in the form of the learning curve per se, the data were subjected to a multivariate analysis in which the number of items recalled on each of the five trials are treated as a mean vector ($df = 5$). The mean vectors are tested for the statistical significance of differences between groups in the nested design.

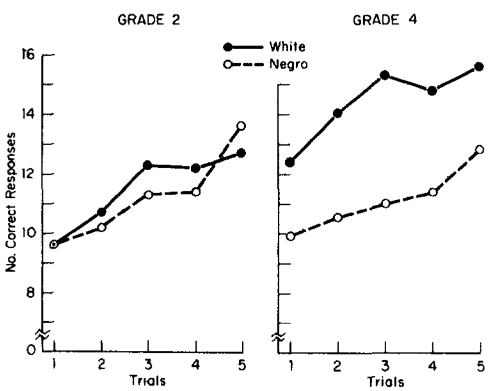

FIG. 2. Amount of free recall of random, categorized 20-item list (5 items in each of 4 categories).

One can distinguish between *strong* and *weak* tests of hypotheses. When a theory involves several hypotheses, any particular study may afford both strong and weak outcomes regarding the hypotheses in question. A strong prediction is one that is essential to the hypothesis and to which the conditions of the experiment allow no extenuating circumstances (other than sampling error). An outcome that is opposite to a strong prediction and is statistically significant, is, of course, the most damaging to the theory. A weak prediction is one that follows from the hypothesis under certain conditions that, in the particular study, are either indeterminate or unsatisfied. That is to say, if the prediction is not borne out, there may be extenuating circumstances (other than sampling error) by the nature of a given study. One seldom controls all factors in any one study, but the amount and nature of the controls in any particular study determine the strength or weakness of the specific hypotheses being tested. So we shall at times make a distinction between the strong and the weak tests of the hypotheses in the present study.

Amount of Recall

The outcomes for amount of recall are shown in Figures 1, 2, and 3. All of the main effects—treatments (uncategorized, categorized, and blocked), race, and grade —are significant beyond the .05 level. But this is of little interest. The relevant points are the race and grade comparisons within the various treatments.

In the uncategorized list (see Figure 1), the Negro and white groups do not differ significantly in either grade (in Grade 2, $F < 1$; in Grade 4, $F = 1.62$, $df = 5/104$, $p > .15$). The analysis shows, in other words, that the variation between the racial groups is very small in relation to the variation within the groups. (It does not, nor could it ever, affirm the null hypothesis.) This finding accords with the hypothesis that the uncategorized list, being mainly a Level I learning task, should show little difference between Negro and white or lower- and middle-socioeconomic-status groups.

The overall difference between Grades 2 and 4, in Figure 1, is significant ($F = 4.57$, $df = 5/104$, $p < .001$). This is consistent with growth in Level I ability during this age span.

In the categorized list (see Figure 2), the white versus Negro difference is not significant in Grade 2 ($F = 1.32$, $df = 5/104$, $p > .25$), but it is significant in Grade 4 ($F = 3.03$, $df = 5/104$, $p < .014$). This accords with the hypothesis that Level II ability, as evoked here by the categorized list, is a larger source of variance, relative to Level I ability, in older than in younger children, and therefore the white–Negro difference is magnified in the older group (Grade 4). However, it is puzzling that Grades 2 and 4 do not differ significantly overall ($F = 1.54$, $df = 5/104$, $p > .18$) on the categorized list, since one should expect age growth in both Levels I and II. The white group shows the expected increase in performance level from Grades 2 to 4, but the Negro groups show no difference whatever. Nor is their recall of the categorized list at all distinguishable at Grade 4 from their recall of the uncategorized list (see Figure 1). This suggests that the Negro groups learn both the categorized and uncategorized lists in the same way, that is, as Level I learning, which accords with the hypothesis but only in a weak sense, since comparisons *between* the uncategorized and categorized lists presume that the lists are equated in difficulty as regards Level I learning. Although we attempted to do this by matching the items in the two kinds of lists on Thorndike-Lorge word frequencies, we cannot be sure how equal the lists really are. If one claims little or no Level II involvement in the fourth-grade Negroes' performance on the categorized list, thereby making it similar to their performance on the uncategorized list, which presumably involves mostly Level I, then one is at a loss to explain the relatively better recall of the categorized list, as compared with the uncategorized, by the second-grade Negroes. Is it because the categorized list is easier in a Level I sense, or is it because some of the Negro second graders learn the list by a Level II process, which would improve recall through clustering? The latter seems unlikely, since, as can be seen in Figure 4, the second-grade Negroes show no more clustering than the fourth graders. For every type of list, we see that the white group shows the larger grade difference, suggesting a steeper growth curve for Level I and especially Level II ability in this age range. We know of no developmental theory or of any previous findings that would lead one to expect no age difference in *any* mental abilities within this age range. The lack of any age difference in recall on the categorized list by the Negro group is therefore inconsistent with expectation and with the previously cited studies by Glasman (1968) and Gerdes (1971). What is highly consistent with the theoretical expectation and with the previous findings is the very small and nonsignificant white–Negro difference on the uncategorized list, the small difference on the categorized list in the younger age group, and the large difference on the categorized list in the older age group.

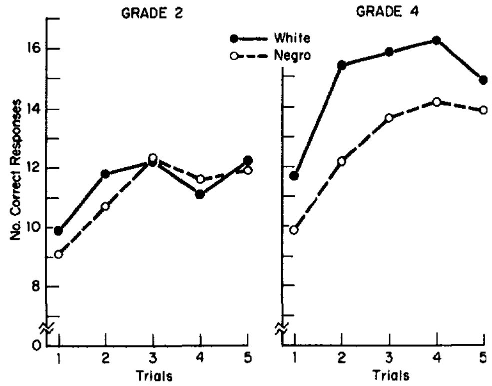

FIG. 3. Amount of free recall of blocked-categorized list (5 items in each of 4 categories) with randomization of items within blocks and order of blocks from trial to trial.

In the *blocked* list (see Figure 3), Grades 2 and 4 differ significantly ($F = 3.91$, $df = 5/104$, $p < .003$). The two socioeconomic-status groups (white vs. Negro), however, do not differ significantly (in Grade 2, $F <$

1; in Grade 4, $F = 1.42$, $df = 5/104$, $p > .20$). There was no prior hypothesis concerning the blocked condition, which was included to find out if making the categories more obvious by blocking would facilitate clustering and recall in subjects for whom a random categorized list does not seem to evoke Level II processes. Blocking, as compared with the random categorized list, had simply no effect on either group at Grade 2; while at Grade 4, it appears that both the white and Negro groups were facilitated by blocking, but the Negro slightly more so than the white, although the effect falls short of significance. Of course, even if subjects acquired the blocked list strictly by rote learning and then recalled the items in much the same order as the order of presentation, they would thereby obtain high clustering scores. But this would not be true Level II clustering and therefore should not facilitate recall.

Category Clustering

Clustering in free recall of the categorized lists was measured in terms of a clustering index, Z, which was devised by Frankel and Cole (1971). It is an improvement over other commonly used measures of clustering, all of which present certain problems that are nicely overcome by the Z index.[3] The Z index is based on the statistical properties of runs. A run is defined as the number of items from the same category that are recalled successively. The length of each run is the number of successive items from the same category. Single items are regarded as runs of one. The expected mean (EM_r) and variance (EV_r) for the number of runs in a randomly selected list of arbitrary length N and number of categories C can be statistically computed (Wallis & Roberts, 1956, p. 571). The Z index of clustering is

$$Z = \frac{EM_r - O_r}{\sqrt{EV_r}},$$

where O_r is the observed runs; EM_r is the expected mean runs in a random series of the same length (N) and number of categories (C) as the observed recall series; and $\sqrt{EV_r}$ is the expected standard deviation of runs in a random series with the same N and C as the observed series.

The Z is thus a standard score referable to the table of the normal distribution for its probability of occurrence. Clustering is defined as the presence of significantly too few runs, that is, fewer runs than would occur in a random output of the same items. As can be seen from the formula, larger Z scores indicate a greater degree of clustering. The Z is a pure measure of clustering, independent of the amount recalled.

Figures 4 and 5 show the group results for the cluster Z scores of the categorized and blocked conditions. Since the expected Z scores is zero for strictly random output, it can be seen that all of the groups showed

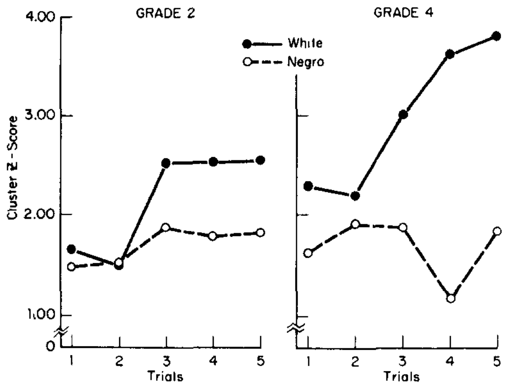

Fig. 4. Amount of clustering in free recall of random, categorized list.

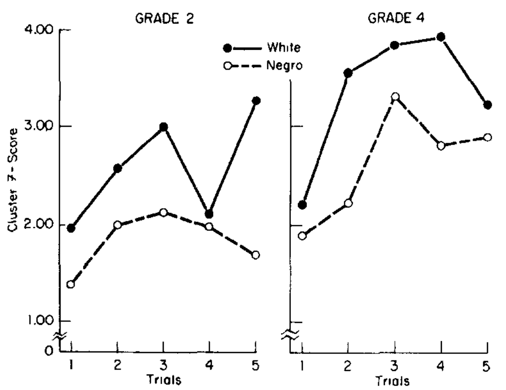

Fig. 5. Amount of clustering in free recall of blocked-categorized list.

[3] We are indebted to Michael Cole for obtaining all the cluster Z scores from our data by means of the computer program he has devised for this purpose.

more than the chance amount of clustering. The method of statistical analysis of the cluster Z scores is the same as that used for the recall data.

The overall main effects of race and grade are significant beyond the .01 level; the main effect of lists (categorized vs. blocked) is significant beyond the .03 level.

Since there were no strong prior expectations concerning the blocked condition, only the amount of clustering in the *random, categorized* list (see Figure 4) provides a strong test of the hypothesis. In Grade 2, the white and Negro groups do not differ significantly ($F < 1$) in amount of clustering, although the difference is in the expected direction. In Grade 4, on the other hand, the white and Negro groups differ markedly ($F = 3.75$, $df = 5/68$, $p < .005$). The nonsignificance ($F < 1$) of the difference between grades in Figure 4 appears to be due largely to the conspicuous lack of any tendency for the fourth-grade Negro sample to show more clustering than the second graders. These results are in accord with the hypothesis that Level II processes (in this case clustering) are more strongly evinced in white (or middle-socioeconomic-status) than in Negro (or low-socioeconomic-status) groups and that this difference in Level II ability is more in evidence in older than in younger children.

In the *blocked* condition (see Figure 5), the difference between grades is significant ($F = 2.91$, $df = 5/68$, $p < .02$). Although in both grades the white group shows more clustering than the Negro group, in neither grade is the difference significant (in Grade 2, $F = 1.40$, $df = 5/68$, $p > .20$; in Grade 4, $F = 1.52$, $df = 5/68$, $p > .19$). No prediction was made for the blocked condition. But it is interesting to note that blocking increased clustering in all the groups and quite markedly in Grade 4. How much of this increase is true clustering and how much is attributable merely to similarity between order of recall and order of presentation cannot be clearly determined in this experiment. But the fact that there is an increase in amount of recall, especially for the Negro fourth graders, suggests some facilitative effect of true clustering, which may have been evoked by the blocking in at least a few of the subjects.

Summary

From Jensen's two-level theory of abilities and from the socioeconomic status and white–Negro differences in these abilities found in previous studies, it was predicted that the Negro (or low-socioeconomic-status) and white (or middle-socioeconomic-status) groups would differ little, if at all, in the Level I task (recall of uncategorized lists). It was also hypothesized that Level II ability is evinced more prominently in older than in younger children and that therefore a small difference between the white and Negro groups would be found in a younger group (Grade 2) and a larger difference would be found in an older group (Grade 4). This was in fact the outcome for both the amount recalled and the amount of clustering of the categorized lists, which are assumed to involve Level II processes as defined in Jensen's two-level theory. These main aspects of the formulation are supported by the present results. A weak hypothesis with respect to these recall tasks involves the general assumption (not peculiar to the two-level theory) that all mental abilities (including Levels I and II) increase with age throughout childhood, though at possibly different rates. In the present study, this was found to hold for the white group in both the Level I and Level II measures. The Negro groups, however, showed almost negligible (and nonsignificant) differences both in amount of recall and in clustering between Grade 2 and Grade 4. Nothing in the present data is capable of accounting for this peculiarity, and only further replications of the study can determine if it is a significant developmental phenomenon in free-recall ability. We doubt that it is. Judging from other studies, it surely does not seem to be a general characteristic of either Level I or Level II measures, which usually show steady, though different, age increments both in white and Negro groups.

REFERENCES

Bousfield, W. A. The occurrence of clustering in the recall of randomly arranged associates.

Journal of General Psychology, 1953, **49,** 429-433.

Frankel, F., & Cole, M. Measures of category clustering in free recall learning. *Psychological Bulletin,* 1971, **76,** 39-44.

Gerdes, B. The relationship of provided organization to free recall learning in two populations. Unpublished doctoral dissertation, University of California, Berkeley, 1971.

Glasman, L. D. A social class comparison of conceptual processes in children's free recall. Unpublished doctoral dissertation, University of California, Berkeley, 1968.

Guinagh, B. J. An experimental study of basic learning ability and intelligence in low socioeconomic status children. *Child Development,* 1971, **42,** 27-36.

Jensen, A. R. Patterns of mental ability and socioeconomic status. *Proceedings of the National Academy of Sciences of the United States of America,* 1968, **60,** 1339-1337.

Jensen, A. R. How much can we boost IQ and scholastic achievement? *Harvard Educational Review,* 1969, **39,** 1-123.

Jensen, A. R. Hierarchical theories of mental ability. In B. Dockrell (Ed.), *On intelligence.* Toronto: Ontario Institute for Studies in Education, 1970. (a)

Jensen, A. R. Level I and Level II performance in low and middle SES elementary school children. In A. R. Jensen & W. D. Rohwer, Jr., *An experimental analysis of learning abilities in culturally disadvantaged children.* (Final Report, U.S. Office of Economic Opportunity Contract 2404) Washington, D.C.: U.S. Office of Education, July 1970. (b)

Jensen, A. R. The phylogeny and ontogeny of intelligence. *Perspectives in Biology and Medicine,* 1971, **15,** 37-43. (a)

Jensen, A. R. The role of verbal mediation in development. *Journal of Genetic Psychology,* 1971, **118,** 39-70. (b)

Jensen, A. R. *Genetics and education.* New York: Harper & Row, 1973.

Keogh, B. K., & MacMillan, D. L. Effects of motivational and presentation conditions on digit recall of children of differing socioeconomic, racial, and intelligence groups. *American Educational Research Journal,* 1971, **8,** 27-38.

Marascuilo, L. A., & Levin, J. R. Appropriate post hoc comparisons for interaction and nested hypotheses in analysis of variance designs: The elimination of Type IV errors. *American Educational Research Journal,* 1970, **7,** 397-421.

Orn, D. E., & Das, J. P. IQ, socioeconomic status, and short-term memory. *Journal of Educational Psychology,* 1972, **63,** 327-333.

Peterson, R. S. The effects of categorization training and class inclusion training on children's use of categories in free recall. Unpublished doctoral dissertation, University of California, Berkeley, 1972.

Rapier, J. L. Learning abilities of normal and retarded children as a function of social class. *Journal of Educational Psychology,* 1968, **59,** 102-110.

Wallace, B. M. An investigation of the educational implications of Jensen's rationale and techniques for differentiating between primary and cultural mental retardation. Unpublished doctoral dissertation, University of Texas, Austin, 1970.

Wallis, W. A., & Roberts, H. V. *Statistics: A new approach.* New York: Free Press, 1956.

White, S. H. Evidence for a hierarchical arrangement of learning processes. In L. P. Lipsitt & C. C. Spiker (Eds.), *Advances in child development and behavior.* Vol. 2. New York: Academic Press, 1965.

(Received November 13, 1972)