AN EXPERIMENTAL STUDY OF BASIC LEARNING ABILITY AND INTELLIGENCE IN LOW SOCIOECONOMIC POPULATIONS

Thesis for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

BARRY JOSEPH GUINAGH

1969

This is to certify that the

thesis entitled

AN EXPERIMENTAL STUDY OF BASIC LEARNING
ABILITY AND INTELLIGENCE IN LOW SOCIOECONOMIC POPULATIONS

presented by

Barry Joseph Guinagh

has been accepted towards fulfillment
of the requirements for

Ph. D. degree in EDUCATION

Lee S. Shulman
Major professor

Date July 17, 1969

ABSTRACT

AN EXPERIMENTAL STUDY OF BASIC LEARNING ABILITY AND INTELLIGENCE IN LOW SOCIOECONOMIC POPULATIONS

By

Barry Joseph Guinagh

The purpose of this study was to examine a theory proposed by Arthur Jensen to explain the different patterns of ability found in middle and low SES levels. He suggests two constructs, Basic Learning Ability and Intelligence to explain the differences. He believes that while Basic Learning Ability is distributed equally in all SES levels, mean intelligence levels are lower in the low SES groups than in the middle SES groups. He attributes the lower intelligence scores in the low SES groups to genetic rather than environmental causes. He also theorizes that Intelligence is functionally dependent on Basic Learning Ability.

In this study, Basic Learning Ability was measured by a Digit Span test and Intelligence was measured by Raven's Progressive Matrices (RPM). Three different populations of third-grade children were tested: (a) Low SES black children (N=105), (b) low SES white children

(N=84), and (c) middle SES white children (N=79). The correlation between the two tests was calculated for the three different samples.

In the second part of the study (the experimental treatment), only those children who scored low on the RPM participated. The procedure was the same for all levels of SES. However, because there were not enough middle SES white children who scored low on the RPM, the middle SES white children were analyzed separately. Twenty children were selected who scored low on both the RPM and the Digit Span test, and twenty children were selected who scored high on the Digit Span test and low on the RPM. In each group of twenty, ten were randomly selected to receive the experimental treatment, which was 7 one-half hour sessions of training on the concepts involved in the RPM. Five were selected to be in a Hawthorne control group, which received instruction on reading skills. Five remained as a pure control group.

The correlations between the Digit Span test and RPM were: low SES white, .10; low SES black, .22; middle SES white, .34. Jensen hypothesized that the correlations would be below .20 for the low SES groups and between .50 and .70 for the middle SES group. The experimental treatment had different effects on the high and low digit span experimental groups for the two races. In the low SES white sample, both the high and low digit span groups had scores on the RPM posttest significantly greater than their respective control group. In the low SES black group, only the high digit span experimental group had RPM posttest scores larger than its control group. The low digit span group did not gain from the experimental treatment. Results were similar for the RPM retention test, which was given one month after the RPM posttest. Ten middle SES white groups made gains similar to the low and high digit span low SES white sample and the high digit span low SES black sample.

The major finding of this study was that scores on the Raven's Progressive Matrices can be substantially increased as the result of a training program. The success of this study was due to a task analysis approach to the development of the training procedure. This approach demonstrates that when the prerequisite skills are taught, the terminal tasks can be mastered.

AN EXPERIMENTAL STUDY OF BASIC LEARNING ABILITY AND INTELLIGENCE IN LOW SOCIOECONOMIC POPULATIONS

By

Barry Joseph Guinagh

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

College of Education

1969

ACKNOWLEDGMENTS

The author would like to acknowledge the financial assistance of the Florida Educational Research and Development Council for sponsoring this project. Their sponsorship enabled more populations to be included in the study.

Tommy F. Tomlinson, Assistant Superintendent for Alachua County (Florida) Public Schools, was particularly helpful in contacting school personnel. He made introducations in five different schools in the county. Ruth Duncan, Director of Curriculum and Research at P. K. Yonge Laboratory School at the University of Florida, worked with us closely also. In total, we dealt with fifteen teachers in six schools. In all of the schools, we experienced complete cooperation, helpfulness, and interest. Without the teachers' cooperation, this research would not have been possible.

I would also like to thank my hard working coaches: Kathy Teply, Susie Weld, Jackie Dufty, and Becky Guinagh. The latter served not only as coach, but as business manager, team psychiatrist, and cheerleader. I am grateful to a talented wife.

TABLE OF CONTENTS

LIST OF TABLES

Table		Page

Appendix

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The differences in the mean IQ scores between socio-economic status (SES) groups in our society is one of the most consistent findings in descriptive psychological research. Correlational studies of the relationship between SES and IQ find coefficients between .25 and .50 depending on the method of classification of SES (Jensen, 1969a). For Eells, Davis, et.al. (1951), the correlations were between .20 and .43. These findings also hold when children's IQs are correlated with parental occupation (McNemar, 1942; Tyler, 1965; Anastasi, 1964). While the differences in IQ scores between SES levels are accepted, the explanation for these differences is an area of much controversy. There are three major explanations given for the differences: (a) The cultural bias explanation, (b) the environmental deficiency explanation, and (c) the hereditary deficiency explanation.

1. The cultural bias explanation states that the low SES groups are just as "bright" as the middle SES groups, but score lower on intelligence tests because the tests are written by middle-class testers and the low SES children are at a disadvantage. This argument, in its pure form, holds that if a culture-free test could be

developed, there would be no differences in IQ scores between SES levels. As of yet, tests constructed according to culture-free criteria show significant differences between SES groups--differences which again favor the middle SES levels (Jensen, 1968a; Haggard, 1954; Ludlow, 1956). For example, low SES black children perform worse on the relatively culture-free Raven's Progressive Matrices than on the ostensibly culturally biased Stanford-Binet Intelligence Scale (Higgins and Silvers, 1958). It is difficult to explain the discrepancy between SES levels by using only a cultural bias explanation. The differences between SES levels do not arise solely as a result of an invalid measuring instrument.

The next two theories accept the different scores for different SES levels as real. The test scores are real in that the IQ test represents an accurate assessment of an individual's intellectual ability regardless of his SES level. The two views hold that differences in scores are the result of individual variations in ability, not faults in the assessment instrument.

2. The environmental deficiency explanation holds that the cause of the poorer performance of the low SES group is due to lack of experiences with the tasks that make up intelligence tests. The deficit is not just a matter of cultural bias, but is a deficit in cognitive ability caused by environmental deprivation.

3. The hereditary deficiency theory holds that the differences in the scores between SES levels is caused by ability levels that are inherited. Since our society has a large degree of social mobility, those with "better genes" acquire better jobs. Thus an individual's inheritance causes both his intelligence level and his SES level. In the environmental deficiency theory, the SES level causes the IQ level.

This study dealt with only the environmental deficiency and hereditary deficiency theories. The assumption basic to these two theories is that the difference in IQ scores is not an artifact of a culturally biased test, but represents real differences between SES populations on a standard set of tasks having predictive and construct validity.

Most psychologists do not see the disagreement between the hereditary and environmental positions as an all or none problem, but as a matter of emphasis. In principle, intelligence is achieved through an interaction between the environment and an individual's genetic ability. However there is much disagreement as to the relative importance of the environment, the individual, and the nature of the interaction. There is also the question of the stability of intelligence, irrespective of the source of variation attributed to nature and nurture. The most thorough recent review of these problems is that of Jensen

(1969a). His research has been the theoretical foundation on which this study was built. However, while Jensen's position supports the hereditary deficiency explanation, this study attempts to demonstrate the validity of the environmental deficiency position.

Nature of Intelligence

It is appropriate at this point to examine the nature of intelligence. Historically, intelligence tests were first developed to separate the slow students that could not profit from a classroom situation from children that could profit from such a situation. Binet, the father of modern intelligence testing, approached the problem in a pragmatic way. He determined what skills first grade teachers expected from their children and developed a test to measure these abilities. That this operational definition of intelligence is a success in prediciting achievement today can be seen in present day correlations between IQ and achievement. The Stanford-Binet predicts scholastic achievement with an average validity coefficient of about .50 to .60 (Jensen, 1969a; Q. McNemar, 1942).

Is there any reason to have two separate tests if the correlation between intelligence and achievement is so high? Perhaps intelligence tests and achievement tests measure the same content. While there is much similarity between the items on many intelligence tests and achievement tests, intelligence tests which have no content similarity to achievement tests still have a high correlation with school achievement. For example Raven's Progressive Matrices, a test which calls for visual completion of a pattern by picking the correct piece from the alternatives presented, had a correlation of between .30 and .50 with academic achievement (Burke, 1959). The correlation between the Raven's Progressive Matrices and the Stanford achievement reading subtest for different grade levels in a middle SES school were the following: Second (N=70), .53; fourth (N=48), .50; sixth (N=45), .45 (based on the author's preliminary work). Because of the dissimilarity of many of the items in intelligence tests and achievement tests, the distinction between intelligence and achievement is maintained in this study.

What is the evidence that different intelligence tests have anything in common? Perhaps each intelligence test measures a different ability. This does not seem to be the case. Tests which measure seemingly different abilities show a high intercorrelation. For example a vocabulary test shows a correlation of .50 with copying sets of designs with colored blocks (Jensen, 1969a). The common factor found in all these tests of higher mental abilities has been referred to by Charles Spearman as the general factor or "g".

Other psychologists have attempted to break the "g" factor down into component parts. Thurstone (1938) developed through factor analysis, a list of not one, but six primary mental abilities. These six abilities were thought of as the elements which combine to make general intelligence. Thurstone tried to produce six pure scales of these abilities. If the six scales were uncorrelated, then the abilities were distinct and Spearman's "g" would be discredited. However Thurstone found that the scales were intercorrelated and Spearman's theory stood (McNemar, 1964). Burt (1958) says:

> In nearly every factorial study of cognitive ability, the general factor commonly accounts for quite 50% of the variance (rather more in the case of the young child, rather less with older age groups) while each of the minor factors accounts for only 10% or less.... For all practical purposes, almost every psychologist--even former opponents of the concept of general intelligence, like Thorndike, Brown, Thomson, and Thurstone--seems in the end to have come round to much the same conclusion, even though, for theoretical purposes, each tends to reword it in a modified terminology of his own.

The issue is not whether the general factor exists, but what emphasis it should be given when compared with the specific abilities that can be identified through factor analytic procedures. Theorists such as Guilford (1967) hold that too much information is lost when only a global intelligence is measured. This may be true, but a global concept of intelligence adquately met the needs of this study.

Intelligence and the Ability to Learn

What exactly is intelligence? In the discussion of intelligence thus far it has always been related to other constructs. Can it stand alone? Many psychologists feel that a definition of intelligence is a futile exercise. The most important point about intelligence is that is has predictive validity. A word definition of intelligence probably does little to shed light on the problem. Spearman believed that "g" consisted of facility in apprehension of one's own experiences, the eduction of relations, and the eduction of correlates (Cronbach, 1960). Jensen defines intelligence as the ability to find relations and correlates, and the ability to solve problems (Jensen, 1969a).

It should be noted that none of these definitions defines intelligence as the ability to learn. Similarly none of the correlates of intelligence discussed thus far is the ability to learn. It is reasonable to assume that learning and intelligence should be highly related, because intelligence is the product of an interaction between the individual and the environment. This interaction surely must involve learning. However the relationship between learning and intelligence is not clear. Woodrow (1946) interpreted his research in the area of ability to learn as indicating that there was no justification for equating mental test scores with the ability to learn. But Woodrow has been criticized for using unreliable gain scores and poor control of starting scores. Both Stake (1961) and Duncanson (1964) found several learning factors which were correlated with general aptitude measured by intelligence tests. The results were interpreted as meaning that no general learning factor is found other than the general aptitude measured by intelligence tests. Anderson (1967) believed that the results found by Stake and Duncanson are actually very similar to those that Woodrow found; Stake and Duncanson just chose to interpret the findings differently.

An examination of intelligence tests, however, shows that the items generally do not involve new learning. For example, the Wechsler Intelligence Scale for Children (WISC) appears to be tests of problem solving and information. The tests of Information, Comprehension, Arithmetic, Similarities, Vocabulary, Picture Completion, Picture Arrangement, Block Design, Object Assembly, and Mazes all involve problem solving and are dependent on some background experiences. Nothing new is learned by the subject. The other two subtests on the WISC are Digit Span and Coding which involve short term memory learning.

If the scores of intelligence tests represent the interaction of heredity and environment, intelligence tests alone cannot solve the heredity-environment problem. What is needed is a more basic test, one that would not involve complex problem solving, but would involve only simple learning. Serial or paired-associate learning would seem to be closer to genetic ability than problem-solving tasks. Jensen (1968a) has found that the simple learning tasks give unexpected results for the children in the 60 to 80 IQ range from different SES levels. Jensen has noted that children of different SES levels with IQs in the subnormal range seem to be functioning at different levels. Children in the low SES with IQ scores in the 60 to 80 range generally seem much brighter than children from a middle SES population with similar IQs. Can it be that for the middle SES level a low IQ score is caused by a lack of genetic ability; whereas, for the low SES level a low IQ score is caused by either a lack of genetic ability and poor environment or solely a poor environment? A learning test might reveal this difference.

Jensen (1968a) has studied children's learning ability directly by actually having the children learn. The learning tasks studied have been varied: Serial and paired-associate learning (Jensen, 1961; Rapier, 1969), trial-and-error learning (Jensen, 1963), and free recall (Jensen, 1961). All used a variety of materials and methods of presentation. The most recent work used the digit span task. Memory of digit series produced the least differences between low and middle SES groups (Jensen, 1968c). But while the correlation between digit span and the Stanford-Binet IQ scores for the normative population was .75 (corrected for attenuation[1]), the correlation between IQ and digit span scores for low SES children was found by Jensen (1968a, 1968c) to be below .20.

This correlation between intelligence and digit span was not the result of cultural bias in the intelligence tests. When the Raven's Progressive Matrices was used as a culture-free intelligence test, the SES differences were even more exaggerated (Jensen, 1968c). The magnitude of the discrepancy between digit span ability and the Raven's Progressive Matrices' performance when a black ghetto school was compared with a white suburban school was revealing. The 30 lowest scoring children in digit span in the suburban school (the lower 6.1 percent of the children in grades 4, 5, and 6) and the 30 highest scoring children in digit span in the ghetto school (the upper 7.9 percent of grades 4, 5, and 6) were compared. The mean digit span test scores (expressed as percent of maximum score) of the suburban and ghetto groups were 38.7 and 65.3 respectively. Their corresponding Raven's Progressive Matrices' scores (again expressed as percent of maximum possible score) were 72.6 and 64.7 respectively (Jensen, 1968c).

---

[1]Attenuation is a correction procedure for correcting the unreliability of variables being correlated. It is calculated by dividing the obtained coefficient for the correlation by the square root of the product of the reliabilities of the two measures being correlated.

As can be seen, many low SES children show a level of ability on the digit span tests that would be entirely unexpected from their low IQs or their poor scholastic achievement. The children's learning performance, however, often corresponds to the teacher's judgment of the child's brightness when observed in play or in social situations. The differences between the SES levels is shown in the hypothetical scatter-diagrams shown in Figure 1. These patterns show the differences between Basic Learning Abilities and IQ in low SES and upper-middle SES groups (from Jensen, 1969a).

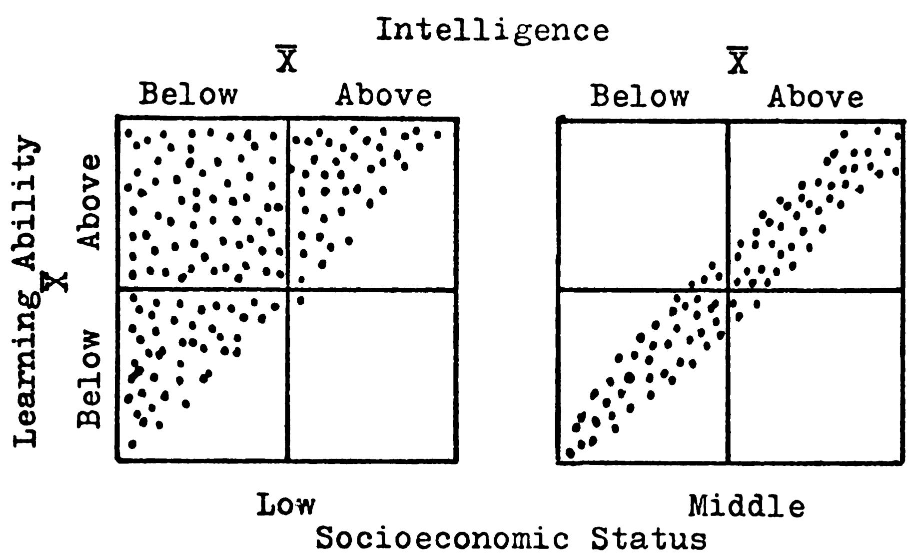


Figure 1. Schematic illustration of the essential form of the correlation scatter-diagram for the relationship between basic learning ability and intelligence in low SES and upper-middle SES groups (Jensen, 1969a).

Jensen's Two Ability Constructs

How can the different patterns in Figure 1 be explained? Jensen (1968c) believes that a minimum of two constructs are needed to explain the data. One is Intelligence which has already been discussed. In order to explain the low SES data, another dimension must be added. Jensen calls the second ability construct Basic Learning Ability (BLA). It is best measured by digit span and rote serial learning. The difference between Intelligence and BLA is in the amount of problem-solving activity required of the testee. For intelligence tests (such as the Raven's *Progressive Matrices*), the testee must bring more covert "mental" activity (discrimination, generalization, verbal mediation, deduction, induction, and hypothesis testing) to bear on the task in order to perform successfully (Jensen, 1968c). These abilities are not required for BLA.

Jensen believes that a score on any test is a phenotypic expression of two separate genotypic sources of variation--Intelligence and BLA. The two abilities are functionally related in that Intelligence is dependent on BLA. Poor BLA is sufficient cause for poor performance in Intelligence, but good BLA is necessary but not sufficient for good performance in Intelligence. Jensen (1968c) believes that the minimal hypotheses needed to explain the patterns of scores in the low and middle SES levels are: (a) Intelligence and BLA are genotypically independent.

(b) Intelligence is functionally dependent on BLA.

(c) While BLA is distributed equally in middle and low SES levels, scores for intelligence tests are lower for low SES groups than for middle SES groups. In the heredity-environment uncertainty question, this theory takes a heredity position.

Alternative Environmental Explanation of the Data

An environmental explanation of the data would hold that the low distribution of intelligence test scores in the low SES level is due to cognitive deprivation. Given experience with the concepts tested on IQ tests, the distribution of the low SES group would become closer to that of the middle SES.[2] If this were the case, another set of minimal hypotheses could explain these findings. The basic hypothesis is that abstract abilities are acquired with experience through BLA. This would explain the middle SES data in which Intelligence and BLA are highly correlated. The middle SES individuals have an adequate environment and can approach the potential their BLA allows. In the low SES groups the low IQ-high BLA individuals exist because of disadvantaged learning environments. If this is the case, it should be possible to increase these low SES students' IQs by giving them

[2]The cultural bias school would also say that the low SES group would improve after practice. However, their usual approach would be to change the test and make it culture free, rather than change the child.

training in the concepts tested by intelligence tests. Thus the low SES and middle SES scatter-diagrams would be similar after training. If the low IQ-high BLA children improve after training and the low IQ-low BLA children do not, it would indicate that Intelligence is functionally dependent on (a) BLA and (b) experience.

Purpose and Hypotheses

The purpose of this study was to examine the increase in IQ scores after training for children of different SES levels and racial groups when BLA levels were taken into account. The experiment was repeated for three populations: low SES black, low SES white, and middle SES white.

The first part of the study attempted to replicate findings reported by Jensen (1969a). These findings of Jensen are reported graphically in the scatter-diagram in Figure 1. A similar scatter-diagram using the data from this study is presented. In the second part of the study, the experimental section, the general expectation was that children with high BLA would improve more on the Raven's Progressive Matrices after training than children with low BLA would improve after identical training. Specific predictions are made in Chapter III.

Summary

Intelligence still remains an elusive construct in psychological research. In Chapter I, three theories explaining why different SES levels have different intelligence scores were discussed. Basically the issue raised by these intellectual differences is the relative importance of heredity and environment in the shaping of intelligence. The nature of intelligence and its relation to other constructs, such as learning, were also discussed. This discussion led to Jensen's two ability constructs, Basic Learning Ability as tested by digit span, and Intelligence as tested by Raven's Progressive Matrices. With these two constructs, the differences Jensen finds between SES levels can be explained. Jensen believes that: (a) The genes determining Basic Learning Ability and Intelligence assort independently. (b) Intelligence is functionally dependent on Basic Learning Ability. (c) While Basic Learning Ability is distributed equally in middle and low SES, scores for intelligence tests are lower for low SES than middle SES.

The alternative explanation tested in this study is an environmental position. It holds that the potential for both Intelligence and Basic Learning Ability is distributed equally in both middle and low SES. The differences between the two SES levels are caused by environmental differences.

The study had two parts: First, the relationship between Intelligence and Basic Learning Ability upon which Jensen bases his theory was examined (see Figure 1). Second, children scoring low on the Raven's Progressive Matrices were trained on the concepts involved in this test to see if their scores could be raised.

CHAPTER II

REVIEW OF LITERATURE

The first studies reviewed examine the relationship between Intelligence (IQ) and Basic Learning Ability (BLA) for different socioeconomic status (SES) levels. Next, research dealing with attempts to change IQ scores with training are reviewed.

Correlation Between IQ and BLA by SES Levels

Jensen (1961) compared learning abilities of Mexican-American and Anglo-American children in the fourth and sixth grades. Although the samples were small (18 for each SES), Jensen found that the distribution of BLA was not substantially different between the two groups. The learning tasks consisted of immediate recall of objects, serial learning, and paired-associate learning of familiar and abstract objects. The serial learning test had the highest reliability for the learning tests. Jensen found in this study that Anglo-American children of low IQ show a lower Basic Learning Ability than Mexican-Americans of the same IQ level.

In another study by Jensen (1963), three IQ groups were tested on BLA. The three groups were classified as retarded (IQs between 50 and 75), average (IQs between

90 and 110), and gifted (IQs above 135). The task consisted of learning by trial and error to associate different stimuli with different responses. Although the learning test correlated with the IQ score in all groups (as low as .19 in the retarded group), Jensen found that four of the thirty-six retarded children learned faster than the average gifted child. These retarded children had not been misclassified by the IQ test, because the IQ test was much more indicative of the scholastic achievement of these children than the learning test.

Rapier (1968) compared the BLA of normal and retarded elementary school children from middle and low SES backgrounds on a series of paired-associate and serial learning tasks. All the children were white. She found the BLA of the retardates from low SES to be greater than that of the retardates from the middle SES groups. She also found that IQ scores were more highly correlated with BLA for the middle SES group (.43) than for the low SES group (.22).

In another study by Rohwer (1967), reported by Jensen (1968c), children, ages 4 to 6, were tested on the Peabody Picture Vocabulary Test and on paired-associate learning. The correlation between mental age and learning scores (with chronological age partialed out) was .10 in the low SES group (N=100) and .51 in the middle SES group (N=100). In this study the low SES children were black and the middle SES children were white. Although these groups differed in IQ by 18 points, they showed no significant differences in serial learning. In serial learning the correlations with mental age (chronological age partialed out) were .10 and .36 for the low SES and middle SES groups, respectively.

In another study Jensen (1968c) studied children in the fourth and sixth grades in an all black school in a low SES neighborhood and an all white school in an upper-middle class suburban neighborhood. The correlation between digit span and the Raven's Progressive Matrices was .33 in the low SES group (N=60) and .73 in the upper-middle SES group (N=60).

In summary, the correlations found between various intelligence tests and various learning tests were .43, .51, .36, and .73 for the middle SES level and .22, .10, .10, and .33 for the low SES level. Jensen hypothesized that the form of the scatter-diagrams in Figure 1 accounts for the differences in correlations between the SES levels.

Changes in Mean IQ Levels

The work done attempting to change IQ levels is vast. This is one of the central objectives of compensatory education. However, this study does not review the literature related to compensatory education because of the many complex variables present in such large scale field research. Jensen has discussed in detail what he describes as compensatory education's failure to raise the IQs of a deprived population (1969a). Other reviews have found essentially the same disappointing results (U. S. Commission on Civil Rights, 1967; Bereiter and Engleman, 1966; Hodges and Spicker, 1967). The present objectives of this study were more circumscribed than those of compensatory education. The experimental treatment in this study is better described as coaching. The education which the subjects received was centered around the concepts covered in the test; the test does not sample from a broader educational experience.

The Effects of Coaching on IQ Levels

A number of investigators have attempted to increase examinees' scores on intelligence or aptitude tests by means of special training or coaching. The findings, as reviewed by Anastasi (1964), show that, in general, coaching produces significant gains in mean scores.

Greene (1928) used two types of training procedures to investigate effects of coaching on Stanford-Binet scores for seven year olds. One group (N=29) was directly coached on the material in the Stanford-Binet. A second group (N=33) was coached with material similar but not identical to the Stanford-Binet. The coaching for both groups lasted only two hours. The groups were retested three weeks, three months, one year, and three years after the pretest. The gains for the directly coached group were

28 points at three weeks, 17 points at three months, 12 points at one year, and 4 points at three years. The gain for the group using similar but not identical materials were 8 points at three weeks, 12 points at three months, 5 points at one year, and 5 points at three years. Anastasi (1964) attributed the decline partly to forgetting and partly to the variation in items on the Stanford-Binet at different age levels. In the one and three year retests, the children were being tested on tasks unlike those on which they had been coached.

Vernon (1954) discussed the effects of coaching versus test-taking practice on English children who were ten years old. The practice group worked eight intelligence tests. The coaching group worked with a teacher in a question and answer session using a commercial book of intelligence test questions. Both the control and coached groups took the pretest and posttest. The over-all mean IQ gain for each group was: control group, 4.7 points; practice group, 11.1 points; coaching group, 6.4 points.

Both the practice and coaching methods seemed to have advantages and disadvantages. While the practice groups showed the greatest mean gains, the individual gains were dependent upon the initial IQ level. At an IQ of 80 the gain was 4 points; at an IQ of 100, 11 points; at an IQ of 120, 18 points. Without the guidance, the weaker students were less able to gain from the practice.

On the other hand the coaching sessions, although lasting the same length of time as the practice sessions, probably did not cover the amount of material covered in the practice sessions. The author reported that the poorer students seemed to make more improvement than the better students in these coaching sessions. It seems that gains in IQ scores consist of two parts: (a) Improvements due to familiarity with the test format and general test sophistication, and (b) improvements due to increased cognitive ability or capability to complete the specific task in question.

Coaching for the Raven's Progressive Matrices

Jensen stated in the Harvard Educational Review (1969a) that, "I have found no studies that demonstrated gains in relatively noncultural or nonverbal tests like Cattell's Culture Fair Tests and Raven's Progressive Matrices." He attributes much importance to two coaching studies (Jacobs, 1966; Jacobs and Vandeventer, 1968) on the Raven's Coloured Progressive Matrices test, a test for children eight and under using the same basic format as the Raven's Progressive Matrices. In a personal communication, he stated that:

> The information you seek concerning training children on the Progressive Matrices is contained in a number of papers by Dr. Paul Jacobs of the Educational Testing Service, Princeton, N.J. The two most important papers are not published but can be obtained from ETS....This work is a "must" for anyone interested in using the Matrices as a research instrument (6/16/68).

These studies were examined because they represent data for Jensen's statement that intelligence is an inherited ability. The studies did not increase test scores of children after coaching, and thus provide evidence for Jensen's belief that the environment has only a small effect on intellectual ability. Jensen states:

> The relative importance of hereditary and environmental factors for Level II processes [Intelligence] is still obscure, although the high heritability of tests like the *Progressive Matrices* and other good measures of abstract intelligence suggest that Level II functions have a genetic basis as well as being influenced by environmental factors. In brief, it seems unlikely that Level II abilities develop solely out of the interaction of Level I [BLA] abilities and environmental influences (1968c).

In the first study, Jacobs (1966) designed a training procedure to increase the scores of 48 first-grade middle-class children on the Raven's *Coloured Progressive Matrices* test. The test displays a pattern with a missing piece; below the pattern are six pieces to the pattern, one of which would complete the pattern correctly. The training used techniques of programmed instruction. Remedial loops were developed which led the child gradually toward success on six of the thirty-six items on the test. For example, in test item A7 the child had to take into account both vertically and horizontally oriented stripes. The pattern of an early remedial frame for that item contained only vertical stripes, and the child had to choose between a piece containing vertical stripes (the correct answer) and a piece containing horizontal stripes. The child was not allowed to test different answers by moving the piece in place. There was no specific instruction on the other thirty items in the test. The total training session lasted only one-half hour. Immediately following the training, the posttest was given. A retention test was given two months later. The posttest showed a greater than chance improvement on the specific items that were the focus of the training. However the retention test indicated no advantage for the experimental children over the control children. There was no general transfer to the items not covered in the training in either the posttest or the retention test. Jacobs' conclusion was that the carefully designed training procedures produced only a temporary gain for specific items in the Coloured Progressive Matrices test. The emphasis of the study was the programmed training used and the author felt that much had been learned. But a half hour training lesson can hardly be considered more than an orientation to the task. Therefore it is understandable that the gains were meager or nonexistent.

The second experiment Jensen referred to was by Jacobs and Vandeventer (1968). The study was an outgrowth of the previous research. The training procedures were more general, and concepts which were in the test were taught. There were actually two training programs used; one consisted of 59 frames and dealt with holistic easy patterns; the second consisted of 64 frames and dealt with "the apprehension of discrete figures as spatially related wholes." The first training procedure was used unless the child was correct on the first ten items. If that happened, the second training program was used. The child was told whether his choices of answers were correct. However, he was never given any explanation of why a particular alternative was right or wrong. The authors felt that a gradual progression in difficulty of frames would be more effective as a teaching technique than the use of verbalization about the desired behavior. The child also could move the alternatives into the space in the pattern to see how they fit. After the first few training frames, the child was encouraged merely to point to the correct alternative if he was sure of the answer, rather than pick it up and place it in the blank spot. He was always allowed to move the alternative however if it was necessary.

The subjects were 44 first-grade and 39 third-grade boys and girls. The training was completed in one sitting and was immediately followed by the posttest. A retention test followed six weeks after the posttest. The results of the experiment showed that the trained children did no better than the control children on either the posttest or the retention test. Again the major criticism raised against this study is that the training sessions were not of sufficient length to reach the children. For Jensen to say that these studies are a "must" is certainly an overestimate of their value. They are only a preliminary look at training for the Progressive Matrices and cannot be considered conclusive.

Summary

The first studies reviewed dealt with the correlation between Intelligence and Basic Learning Ability. The results, although of the most preliminary nature, seemed to indicate that the correlations found for the middle SES level are higher than those found for the low SES level. Jensen stated in a recent article (1969a) that a sample of 5000 children is presently being tested to determine more exactly the relationship between Intelligence and Basic Learning Ability.

The second area reviewed was the effects of coaching on intelligence test scores. The research by Vernon (1954) indicated that mean gains of up to 11 points were possible after coaching. The gains were caused by improvements due to familiarity with the test format as well as actual improvements in cognitive ability. Several studies were reviewed in detail in which coaching failed to produce improvements on the Raven's *Coloured Progressive Matrices*, a test similar to the Raven's *Progressive Matrices*. The major criticism of these studies was that the total time in training was not of sufficient length to reach the children.

CHAPTER III

DESIGN

Sample

Three different third-grade populations were studied:

1. Low SES black children (N=105) from an all-black school in Gainesville, Florida.

2. Low SES white children (N=84) from two different rural communities near Gainesville. Eleven black children in these schools were eliminated from the study. Seven white children were also eliminated from the study because at least one parent had two or more years of college. For both the low black and low white populations, the educational level of the parents was at most the completion of high school.

3. Middle SES white children (N=79) from the laboratory school at the University of Florida and a middle SES suburban school in Gainesville. Thirteen black students in the classes were eliminated from the study. Fifteen white students whose parents did not have two or more years of college were also eliminated from the analysis. A middle SES black group of children was not tested because it was the author's opinion that there were too few middle SES black children in the community.

Three factors usually make up measurement of SES levels: occupational level, income level, and educational level. In this study the major factor that catogorized the children was the school attended. The different schools served different SES levels in the community. A measure of parent educational level served to make each SES level more homogeneous: Income and occupational levels were not used in the definition of SES level in this study.

Instruments

Two instruments were used in the research: The Raven's Progressive Matrices and a Digit Span test.

The Raven's Progressive Matrices (RPM) was developed in 1938 as a culture-fair test. The test consists of a series of patterns each with a missing piece. The subject is to choose the piece that finishes the pattern from alternatives presented. The test is nonverbal in that reading or language is not directly tested. Raven (1960) reported the test-retest reliability coefficients were approximately .85. Burke (1958) gave a review of reliability studies, and although most coefficients were not as high as those Raven reported (the range of coefficients found for children was from .71 to .88), the coefficients were as high as other standardized intelligence tests.

The correlation between the RPM and other intelligence tests is quite high. For 70 children, ages 9.2 to 10.1, the correlation with the Wechsler Intelligence Scale for Children was .75 for the total score, .69 for the verbal score, and .70 for the performance scale (Barratt, 1956). However, Raven did not believe the test probed all aspects of intelligence (IQ). He spoke of it as a "means of estimating a person's innate eductive ability... a test of a person's capacity to form comparisons, reason by analogy, and develop a logical method of thinking regardless of previously acquired information (Burke, 1958)." Spearman considered RPM a test for measuring "g". While the factor analytic studies show it is not a pure measure of "g", Burke (1958) found a loading of .80 for a "g" factor.

The Digit Span (DS) test was given individually in a format similar to the Digit Span subtest of the WISC. The tester read a series of numbers at one-second intervals. After completing a series, the child attempted to repeat the series orally. The child was only given credit if the series was repeated exactly as given. In the particular test used in this study, the child received as many points as there were numbers in the series. For example, a five digit series was worth five points. The test consisted of 2-two digit series; 2-threes; 5-fours; 4-fives; 4-sixes; 4-sevens; 2-eights. The tester continued to give new digit series to the child until three series were missed in a row.

The repetition of digits is one of the oldest psychological tests used to measure individual differences. Cattell (1948) used a series of letters to test memory as early as 1890. Generally the DS has been considered a test of short term memory. Jensen believes that DS measures a Basic Learning Ability (BLA), an associative ability that requires no self-initiated or problem-solving activity (Jensen, 1968a). No manipulation of the input is necessary. BLA is tapped by tests such as DS memory, serial rote learning, selective trial-and-error learning with reinforcement (feedback) for correct responses, and in a slightly less "pure" form by free recall of visually or verbally presented materials, and paired-associate learning. Jensen's choice of DS is based on his research conclusion that DS shows the smallest differences between SES levels.

Figure 2 shows a two-dimensional plot and the relative theoretical positions of RPM, DS, and other tests. The two dimensions are (a) degree of cultural loading and (b) the conceptual to associative dimension. Both the RPM and DS are culture free, but the two tests are on opposite ends of the associative to conceptual dimension. That is, the RPM is a test of abstract ability, whereas, DS is a test of associative ability.

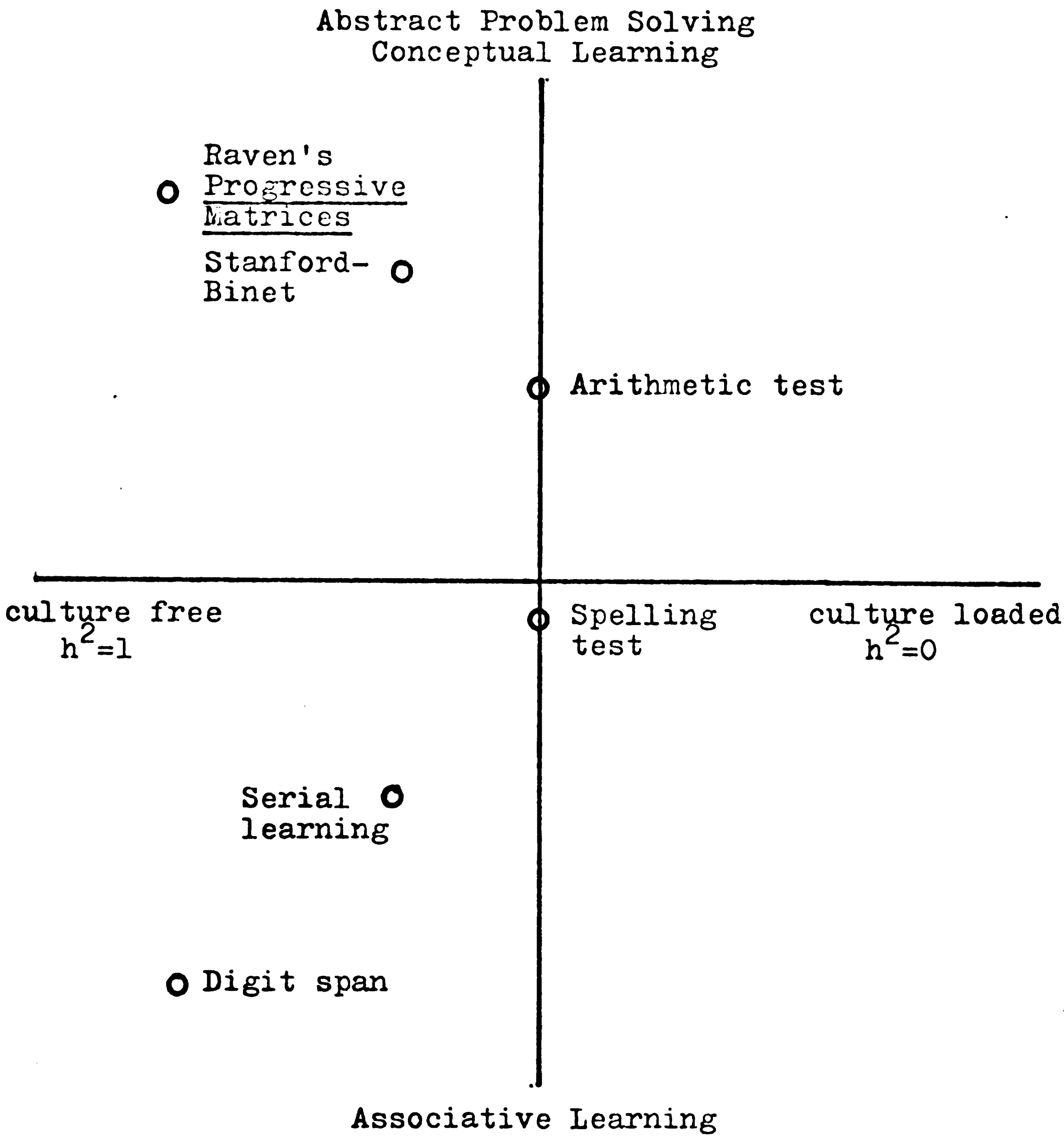


Figure 2. Hypothetical two dimensional plot of the relative theoretical positions of Raven's Progressive Matrices, Digit span, and other tests (from Jensen, 1969a). The degree of cultural loading for a test is based on empirical data of the heretability estimates for the different tests. Notice that this method of estimating cultural loading places no tests in the upper and lower right quadrants, the cultural loaded quadrants.

Procedure

In the first part of the study the relationship between the RPM and DS was examined for different SES levels. Correlations were calculated for each of the three samples. The RPM was given as a group test: The DS was given individually.

For the second part of the study (the experimental treatment), only those children who scored low on the RPM were chosen for the treatment. A score of twenty or less was classified as a low score. Twenty was chosen as the cutoff point on the basis of extrapolation from data collected from second-and fourth-grade children in Haslett, Michigan. The reasoning for this was as follows: Jensen defined an IQ between 100 and 120 as a high IQ and an IQ between 60 and 80 as a low IQ. Since 100 is the mean and 15 points is one standard deviation for IQ tests, the scores for the high IQ children are at the mean and the scores for the low IQ children are about one standard deviation below the mean. The Haslett data for the RPM follows:

Second Grade	mean=26.8	s.d.=9.7	N=70
Fourth Grade	mean=33.8	s.d.=9.5	N=48

The extrapolated mean for third graders based on the above data was 29. One standard deviation below the mean was approximately 20. Therefore scores below 20 should give approximately the same ability range as Jensen

TABLE 1

Experimental Design Replicated
with Low SES White and Low SES Black Samples

Intelligence level	Basic learning ability level			7 half hour training sessions	RPM[a] post-test	one month break	RPM retention test
Low RPM	High digit span	Experimental N=10		on the RPM	X		X
		Control N=10	Hawthorne N=5	on phonic skills	X		X
			Pure N=5	no training	X		X
	Low digit span	Experimental N=10		on the RPM	X		X
		Control N=10	Hawthorne N=5	on phonic skills	X		X
			Pure N=5	no training	X		X

[a]Raven's Progressive Matrices

used. In a few cases, scores higher than twenty were taken to complete the sample when there were not enough scores below twenty in a sample. Seven of the forty low SES black and two of the forty low SES white students had scores on RPM above twenty.

From those students in each sample that made low scores on RPM, twenty were selected with the lowest DS scores and twenty were selected with the highest DS scores. Extreme groups were used so that any differences between the low DS groups and the high DS groups would be more obvious.

The original plan was to have all three populations compared before and after training. However it proved impossible to find enough middle SES white children who scored low on RPM to complete the design. Therefore only in the low SES white and low SES black populations were there enough to do a complete analysis. More discussion of this topic appears in the next chapter.

From each of the groups (low DS-low RPM and high DS-low RPM), ten children were randomly selected to receive the experimental treatment and five were randomly selected for a Hathorne control group. Five remained as pure control and received no treatment.

The experimental group received 7 one-half hour sessions of training on the concepts involved in the RPM. The first two sessions were on a one-to-one basis with the trainer, and, in the five remaining sessions, two children worked with a trainer. The same format was used with the Hawthorne control group. The children in this group received tutoring on word skills through such games as "crazy eights" played with phonic rummy cards by Kenworthy Educational Service, Inc. The Hawthorne group thus received the same contact with the trainer. They were taught verbal skills rather than the skills necessary to solve the RPM. Both the experimental and the Hawthorne groups received their respective treatments within two weeks. At the start of the training, both the experimental and Hawthorne control groups were told that they would be retested on the "puzzle test". With the culmination of the treatment, all three groups, the experimental, the Hawthorne control, and the pure control, were given a posttest on the RPM. After four more weeks with no training, a retention test was given to the same groups that received the posttest.

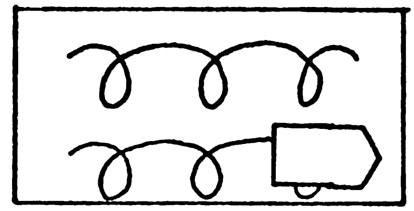
Experimental Treatment

The experimental treatment used a task analysis model as discussed by Robert Gagné in the "Acquisition of Knowledge" (1962). In that article, Gagné discussed the hierarchy of content knowledge necessary for the terminal performance of a task. This hierarchy consists of elements that are a part of the terminal task or subordinate tasks that are logically simpler than tasks higher in the hierarchy. The analysis is based on the content of the material. The subordinate tasks are assumed prerequisites to the terminal task. The hierarchy does not give heuristic strategies to solve the terminal task; it only gives prerequisite knowledge that must logically be present to solve the terminal task.

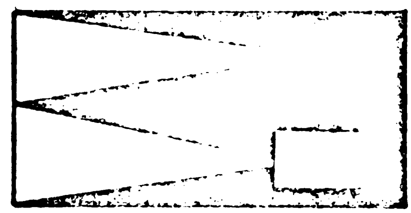
A task analysis of items on RPM was developed by studying the problem-solving tactics used by superior adults and children. The mistakes of children who scored below twenty on RPM were also studied. Based on the observations of these subjects, it appears there are two separate terminal types of figures in RPM. The first is the total Gestalt pattern and the second is a matrix of figures that calls for logical operations to achieve their solutions (i.e., the use of rules). Within the total Gestalt pattern, there are two types of formats that both demand a holistic approach: One format is a continuous pattern (see Figure 3) and the second format is a matrix made up of discrete figures (Figure 4). The two Gestalt-type terminal patterns are called Gestalt-continuous and Gestalt-discrete. The matrix items with discrete figures calling for logical operations are called logical-discrete (see Figure 5).

The Gestalt-continuous patterns (see Figure 3) are the simplest items on the test. Raven describes them as

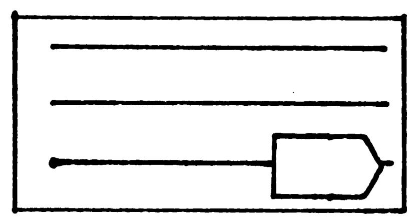
4. The child must be able to judge curved and jagged figures.



3. The child must see differences between a form filled-in and a form which is not filled-in.



2. The child must not be distracted by the whole figure, but must examine the section where the piece is missing.



1. The child must complete continuous figures where the missing piece is the same pattern as the whole figure.

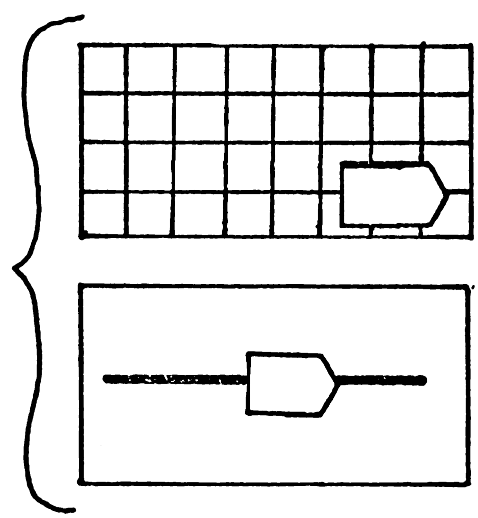
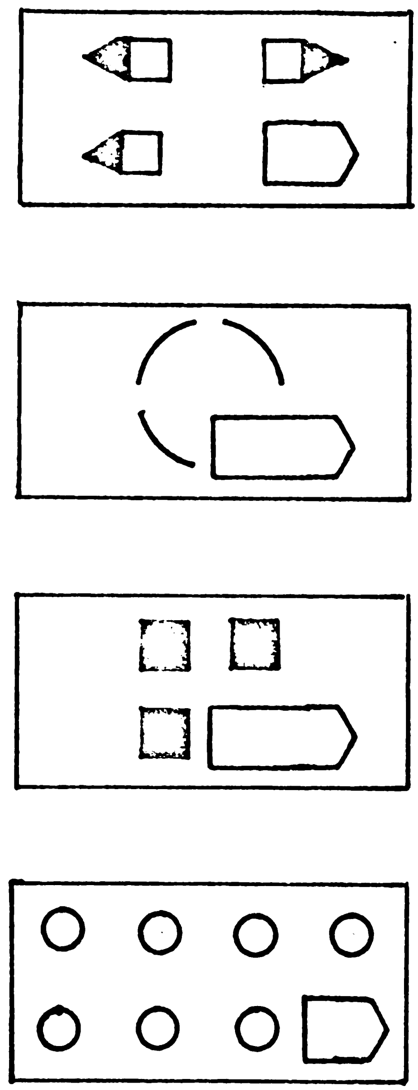


Figure 3. The hierarchy used to teach Gestalt-continuous patterns.



4. The child must be able to complete symmetrical patterns.

3. The child must be able to complete an outlined figure.

2. The child must be able to complete a solid figure.

1. The child must be able to complet figures where missing pieces are the same as the whole figure.

Figure 4. The hierarchy used to teach Gestalt-discrete patterns.

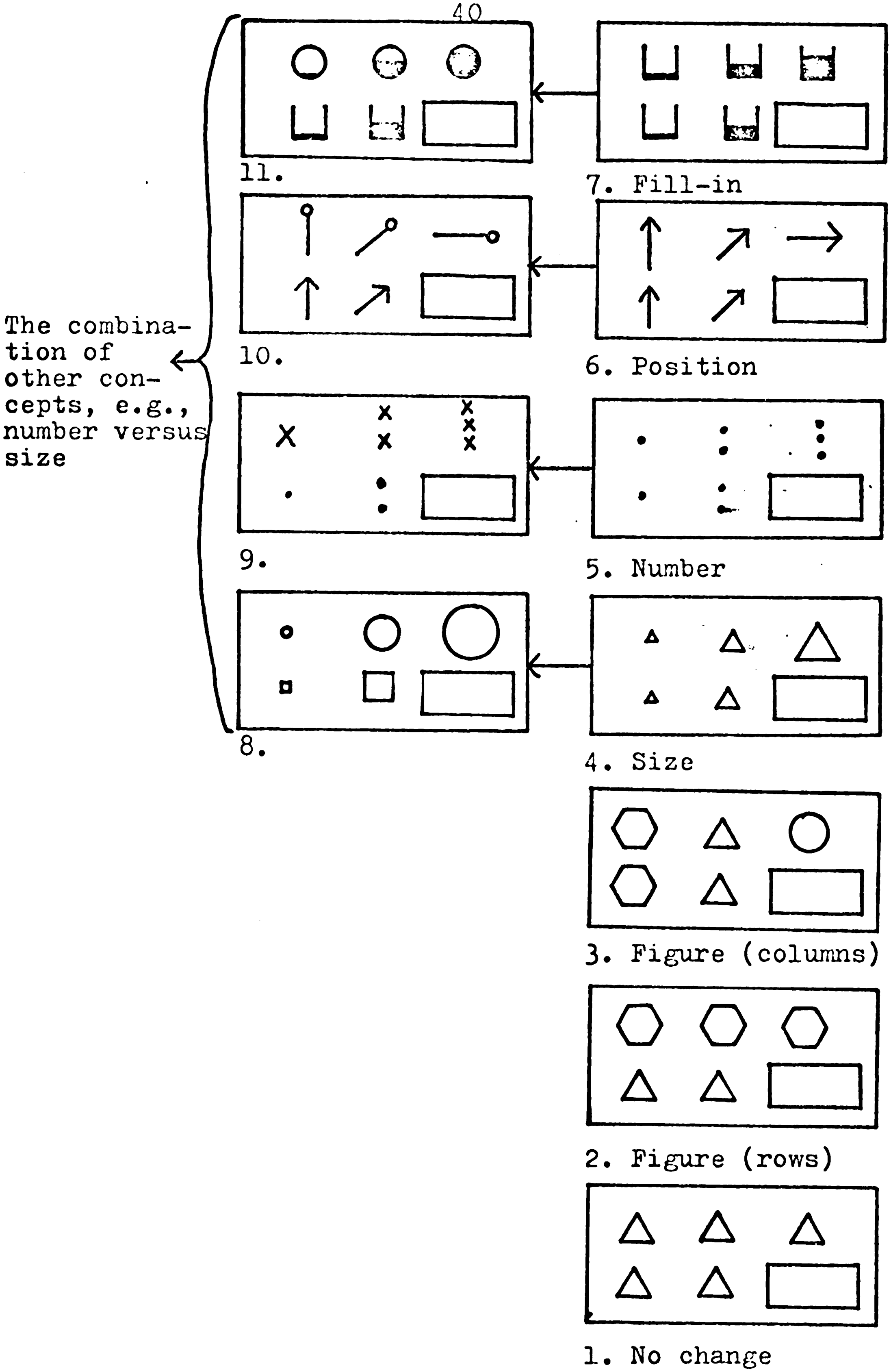


Figure 5. The hierarchy used to teach logical-discrete patterns.

".... simple continuous patterns in which the completion involves perception of differences, similarities and identity of pattern only (Raven, 1956)." In other words, these patterns are not based on logical rules, but are only based on seeing the totality of patterns and judging what the missing piece should be. During training each pattern was presented first with movable puzzle pieces. Each pattern was then presented with fixed alternatives, as in the test. The pattern was usually presented three times, with two, four, and six alternatives.

The Gestalt-discrete patterns (see Figure 4) call for the same intellectual ability as the Gestalt-continuous figures as shown in Figure 3. According to Raven these measure "apprehension of analogous changes in spatially and logically related figures and call for concrete or coherent reasoning by analogy (Raven, 1956). The superior subjects found it most difficult to explain the rational for their solutions because the strategy for solution of the pattern is not based on rules, but is based on seeing a total pattern and viewing the pattern as a whole. Again both movable puzzle pieces and regular alternatives were used.

The logical-discrete patterns (see Figure 5) are discrete matrices, which require the subjects to utilize logical operations. This mental operation is different from the Gestalt-continuous and Gestalt-discrete. Here superior subjects could easily explain exactly what they did to solve the puzzle, because in order to solve the matrix, it was necessary to analyze it into its component parts. Movable alternatives or puzzle pieces were used only if the child needed them. The five major concepts are figure change, size change, number change, position change, and fill-in change. These were learned in both the horizontal and vertical directions. Next figure change in one direction was combined with either size change, number change, position change, or fill-in change in the other direction. Then the other concepts were combined, e.g., number change versus size change. The numbers in Figures 3, 4, and 5 refer to the order of presentation of the tasks.

A general strategy used throughout the experimental treatment was the use of verbal instructions. Most of the experimental treatment was on a one-to-one basis. While the trainer stayed close to the "course outline" and worked his way toward the terminal tasks, the opportunity for the students to ask questions was open, and the trainer was always able to explain a concept in more detail. Jensen has commented that although RPM is a nonverbal test, it is often accompanied by much lip movement. Thus the test has verbal components in it. Gagné (1962) lists four uses of instructions that are important in progressing toward the terminal task. One is the identification through verbalization of the elements present in the task. Second is the identification of the terminal performance. Third is giving verbal cues to aid in recalling previous learning sets, and fourth is using language to guide the subject's thinking.

Figures 3, 4, and 5 give the "course outline" used with the children. None of the training items was from RPM. It was hypothesized that the skills learned from the training sessions would transfer to a posttest of RPM. The Gagné-type hierarchy was modified in order to give the children practice in using multiple choice tests. The general procedure was to give the alternatives to the children as puzzle pieces and let them try out different pieces before they chose the correct one. Next the pattern was presented without the puzzle pieces with less than six alternatives, usually two. Then four alternatives and finally the terminal task was reached with six alternatives for the child to choose from. Since the possibilities for patterns are infinite, only samples of the different levels of the hierarchy were presented.

Hypotheses

Jensen holds that the correlation between BLA and IQ for low SES groups is below .20, and for middle SES groups it is between .50 and .70 (Jensen, 1968a). The first two hypotheses tested these correlations.

1. The correlation between RPM and DS is between 0 and .20 for both the low SES black and the low SES white groups.

2. The correlation between RPM and DS in the middle SES white sample is between .50 and .70.

The next two hypotheses are related to Jensen's thesis that high BLA is necessary for high IQ, i.e., IQ is functionally dependent on BLA. This study attempted to show that training plus a high BLA can affect IQ scores.

3. The means of the experimental groups are significantly higher than the means of the control groups (those not receiving the experimental treatment; i.e., Hawthorne and pure control combined) on the RPM posttest and RPM retention test.

4. The means of the high DS groups are significantly higher than the means of the low DS groups.

The next two hypotheses dealt with the different samples and different experimental treatment.

5. There is no significant difference between the means of the low SES white and the low SES black groups on RPM.

6. There is no significant difference between the means of the pure control and Hawthorne control groups on RPM.

Analysis

The data were analyzed by a 2 X 2 X 2 X 3 analysis of variance using RPM pretest, RPM posttest, and RPM retention test as repeated measures. The first three factors consisted of two levels: Low SES black and low SES white race, high and low DS, experimental and control treatment (control children did not receive the experimental treatment, i.e., the pure control and Hawthorne control groups). The data from the RPM pretest, the RPM posttest, and the RPM retention test were each analyzed by a 2 X 2 X 2 analysis of variance design. The Hawthorne and pure control groups were compared in a separate analysis.

TABLE 2

2 X 2 X 2 X 3 Analysis of Variance
for Repeated Measures

Race	Digit span	Treatment	RPM pretest	RPM posttest	RPM retention test
Low SES black	High digit span	Experimental	N=10	N=10	N=10
		Control	N=10	N=10	N=10
	Low digit span	Experimental	N=10	N=10	N=10
		Control	N=10	N=10	N=10
Low SES white	High digit span	Experimental	N=10	N=10	N=10
		Control	N=10	N=10	N=10
	Low digit span	Experimental	N=10	N=10	N=10
		Control	N=10	N=10	N=10

Summary

Three different populations of third-grade children were tested in this study: (a) Low SES black children (N=105), (b) low SES white children (N=84), and (c) middle SES white children (N=79). All the children were given the Raven's *Progressive Matrices* and a Digit Span test. The correlation between the two tests was calculated for the three different samples.

In the second part of the study (the experimental treatment), only those children who scored low on the *Progressive Matrices* participated. The procedure was the same for all levels of SES. However, because there were not enough middle SES whites who scored low on the *Progressive Matrices*, the middle SES white children were analyzed separately. Twenty children were selected who scored low on both the *Progressive Matrices* and the Digit Span test, and twenty children were selected who scored high on the Digit Span test and low on the *Progressive Matrices*. In each group of twenty, ten were randomly selected to receive the experimental treatment, which was 7 one-half hour sessions of training on the concepts in-involved in the *Progressive Matrices*. Five were selected to be in a Hawthorne control group, which received instruction on reading skills. Five remained as a pure control group.

A 2 X 2 X 2 X 3 repeated measures design was used to analyze the data. Each factor was made up of two levels: Low SES black and low SES white race; high and low digit span; and experimental and control (Hawthorne and pure combined) treatment. The three repeated measures were the *Progressive Matrices'* pretest, posttest, and retention test. A separate analysis was made between the Hawthorne and pure control groups.

CHAPTER IV

ANALYSIS OF RESULTS

The six hypotheses presented in the last chapter are examined in order.

Hypothesis 1. The correlation between the Raven's Progressive Matrices (RPM) and the Digit Span test (DS) is between 0 and .20 for both the low SES black and low SES white samples.

Hypothesis 2. The correlation between RPM and DS in the middle SES white sample is between .50 and .70.

TABLE 3

Means, Standard Deviations, and Correlations Between DS and RPM Pretest

	DS		RPM		Correlations between DS and RPM	
	Mean	s.d.	Mean	s.d.		
Low SES black N=105	40.7	16.0	13.9	4.9	.23	.29*
Low SES white N=84	47.5	18.2	24.6	8.4	.10	.13*
Middle SES white N=79	49.3	13.6	34.1	8.4	.34	.43*

*Correlations corrected for attenuation assuming both measures had a reliability of .80.

Although there was a trend in the correlation that was in the expected direction, there was no significant difference between any of the correlations between DS and RPM pretest (see Table 3). Since Jensen's theoretical correlations were corrected for attenuation, the obtained values were also corrected for attenuation. A reliability coefficient of .80 was assumed for both the RPM and the DS test. After the correction for attenuation, the low SES black correlation was .29, the low SES white correlation was .13, and the middle SES white correlation was .43. The 95 percent confidence range was between .11 and .29 for the low SES black and between -.08 and .34 for the low SES white. Both these ranges are not inconsistent with the correlation of between 0 and .20 Jensen found for low SES populations. The 95 percent confidence range for the middle SES white was between .23 and .59. This finding is not inconsistent with the .50 to .70 correlations Jensen found for the middle SES populations.

Therefore the correlations did not demonstrate or repudiate the validity of Jensen's theory. The actual correlations were not significantly different, but the attenuated correlations could have come from the range that Jensen would have predicted. Jensen does much of his theorizing with attenuated correlations. This makes comparisons difficult because of differences in the reliability of the measures. Attenuated correlations also inflate the difference between correlations, as can be seen in Table 3.

The scatter-diagrams summarizing the relationship between RPM and DS for the three samples are shown in Figures 6, 7, and 8. According to Jensen's theory, a high BLA is necessary for a high IQ score. Therefore if the scatter-diagrams for each sample are divided into quadrants at the means of the DS and RPM scores, the low DS-high RPM quadrant should be empty.

Figures 6, 7, and 8 show that this quadrant was filled in each scatter-diagram. In the low SES black scatter-diagram (Figure 6), 23 of the 105 children were in the low DS-high RPM quadrant. In the low SES white scatter-diagram (Figure 7), 22 of the 84 children were in this quadrant; and in the middle SES white scatter-diagram (Figure 8), 12 of the 79 were in this quadrant. When a grand mean is used as the divider for the quadrants, 45 of the 268 children are in the low DS-high RPM quadrant. The scatter-diagrams give no evidence for Jensen's hypothesis that high BLA is necessary for high IQ. Nor is any absolute DS cutoff point obvious below which a

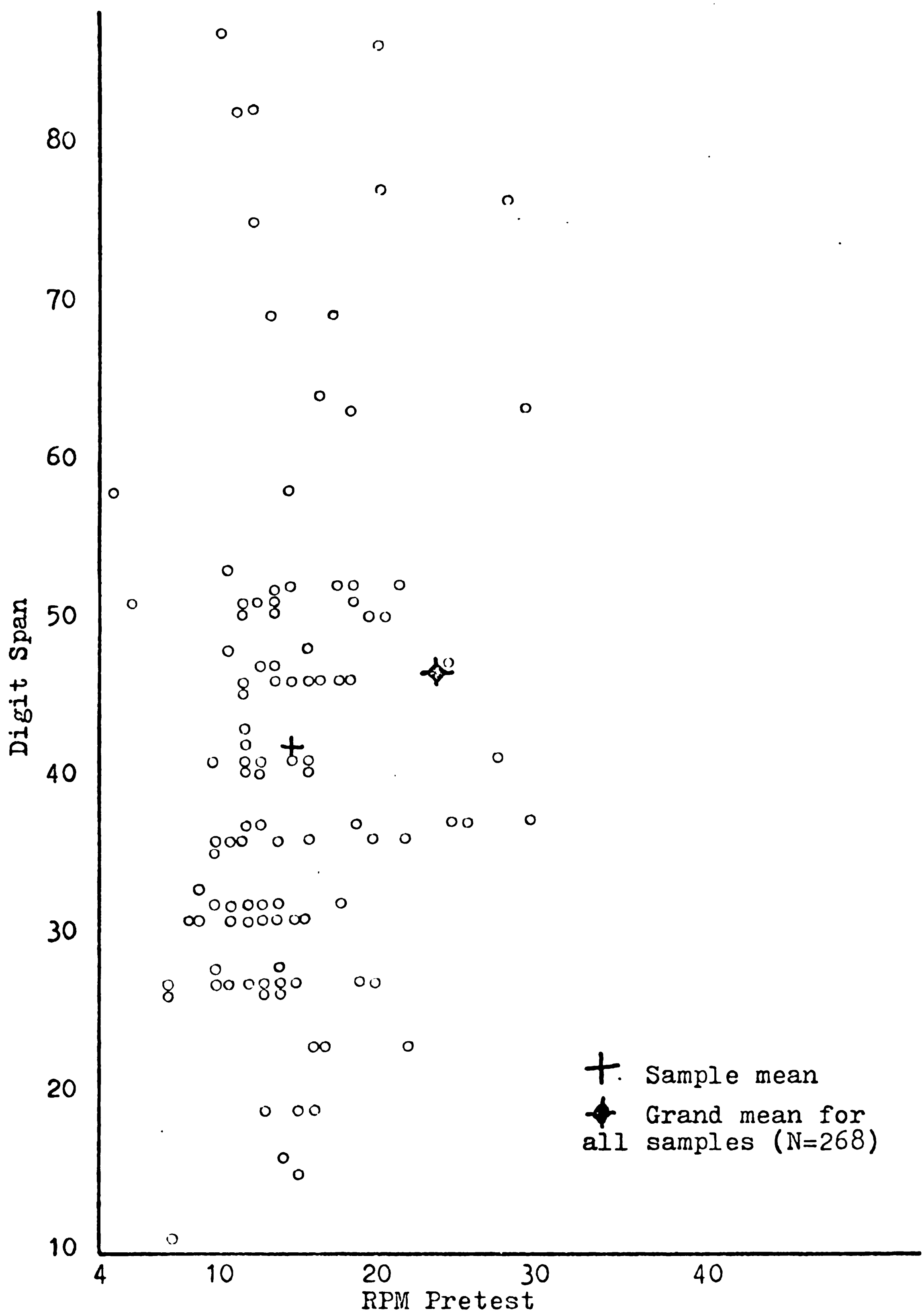


Figure 6. Scatter-diagram for low SES black children (N=105). Correlation between DS and RPM Pretest is .23.

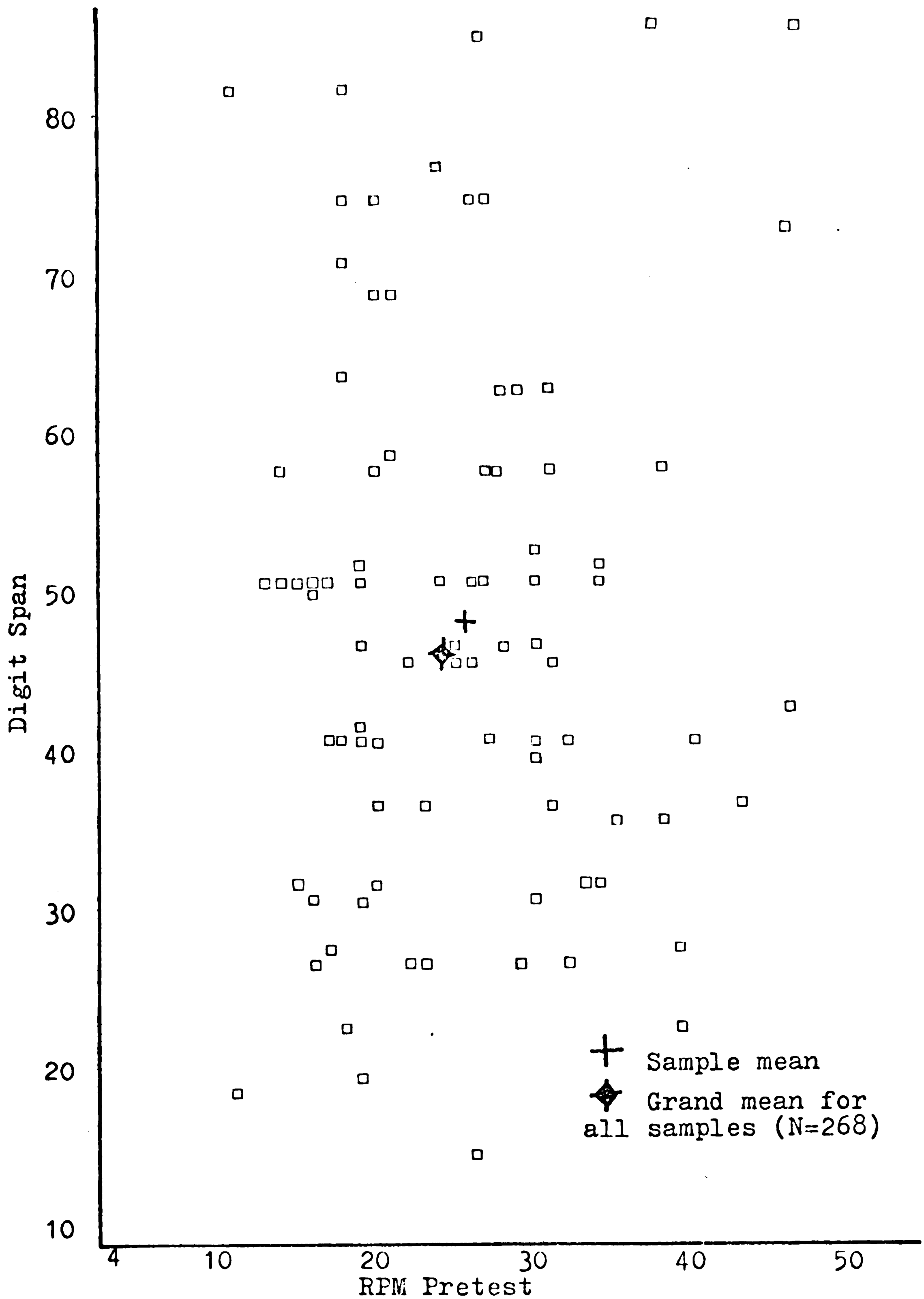


Figure 7. Scatter-diagram for low SES white children (N=84). The correlation between DS and RPM Pretest is .10.

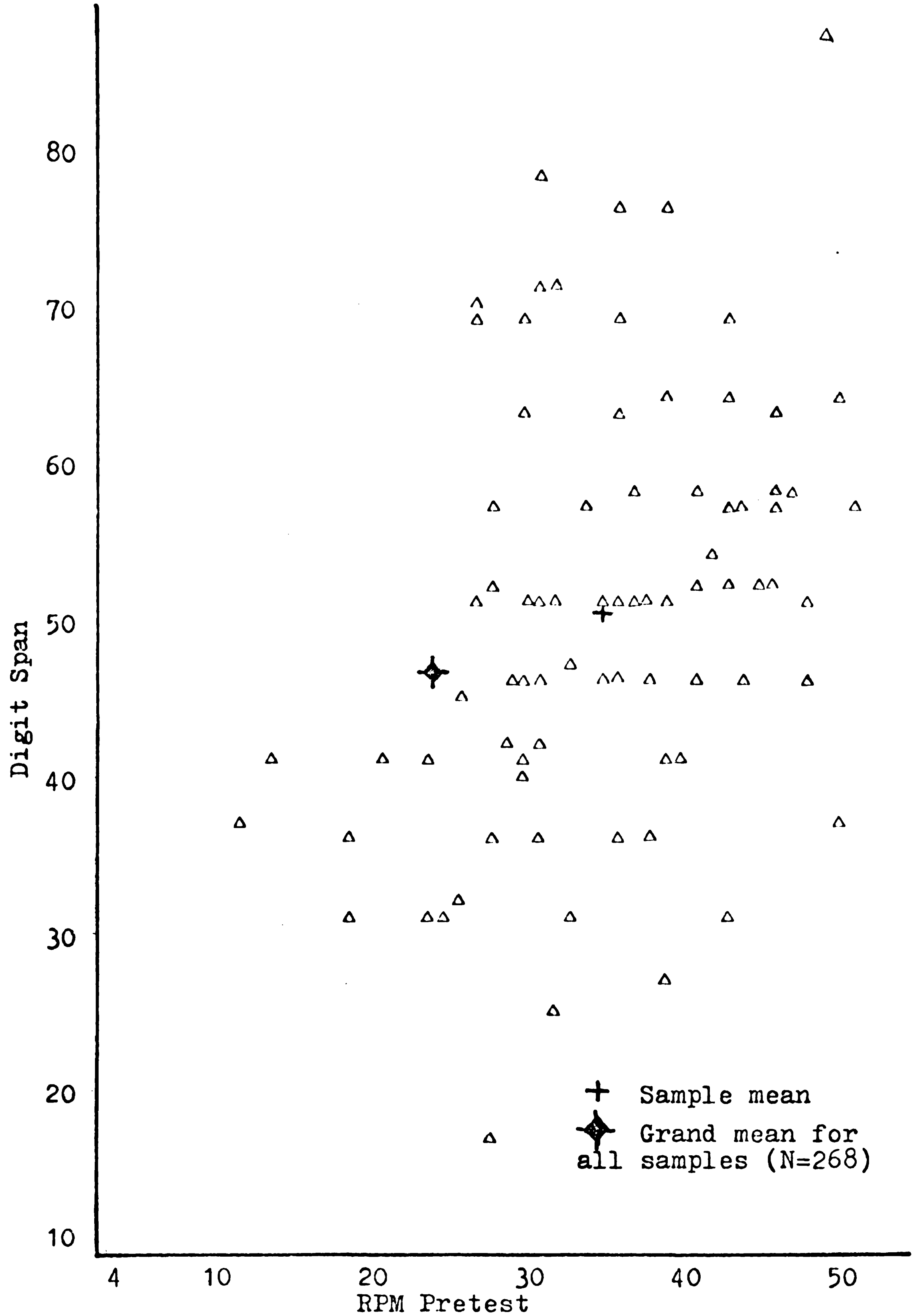


Figure 8. Scatter-diagram for middle SES white children (N=79). The correlation between DS and RPM Pretest is .34.

high RPM score is impossible.

Results of the Experimental Treatment

Table 4 reports the means and standard deviations of the RPM pretest, RPM posttest, RPM retention test, and DS for the experimental and control groups. The results are presented graphically in Figure 9.

Hypothesis 3. The means of the experimental groups are significantly higher than the means of the control groups (those not receiving the experimental treatment, i.e., the Hawthorne control and pure control groups combined). Table 5 reports the analysis of variance using the RPM pretest, RPM posttest, and RPM retention test as repeated measures. The main effect due to treatment for this analysis was significant ($p<.001$).

To make multiple comparisons between the means of the different groups when the over-all analysis of variance was significant, a Duncan multiple range test was used. The comparison between the means of the experimental groups and the corresponding control group using the RPM posttest means show that the means of all the experimental groups were higher than the control groups ($p<.01$) with the exception of the low DS black group which was only 4.4 points above its control group (see Table 6). The RPM retention test results are reported in Table 7. Only in the high DS black and the low DS white groups were the

TABLE 4

Means and Standard Deviations for DS, RPM Pretest, RPM Posttest, and RPM Retention Test

	DS	RPM pretest	RPM posttest	RPM retention test
Black-high DS experimental	64.0 (12.63)	14.7 (2.90)	32.0 (4.27)	33.4 (4.36)
Black-high DS control	63.7 (13.52)	14.2 (2.96)	12.9 (4.85)	16.3 (8.08)
Black-low DS experimental	30.2 (14.95)	14.3 (3.74)	20.8 (6.45)	22.7 (8.89)
Black-low DS control	30.8 (21.34)	16.2 (2.93)	16.4 (6.20)	19.8 (9.16)
White-high DS experimental	64.2 (12.81)	17.7 (3.61)	29.7 (4.29)	28.5 (5.50)
White-high DS control	61.3 (13.06)	18.0 (3.13)	22.2 (5.84)	21.3 (7.35)
White-low DS experimental	28.1 (18.30)	16.5 (3.30)	32.3 (5.10)	31.7 (6.86)
White-low DS control	32.2 (9.04)	18.8 (3.06)	20.5 (4.57)	18.8 (8.92)

Note. Mean scores are followed by the standard deviation in parentheses. N=10 in each cell. The same subjects were used in the DS, RPM pretest, RPM posttest, and RPM retention test in each row.

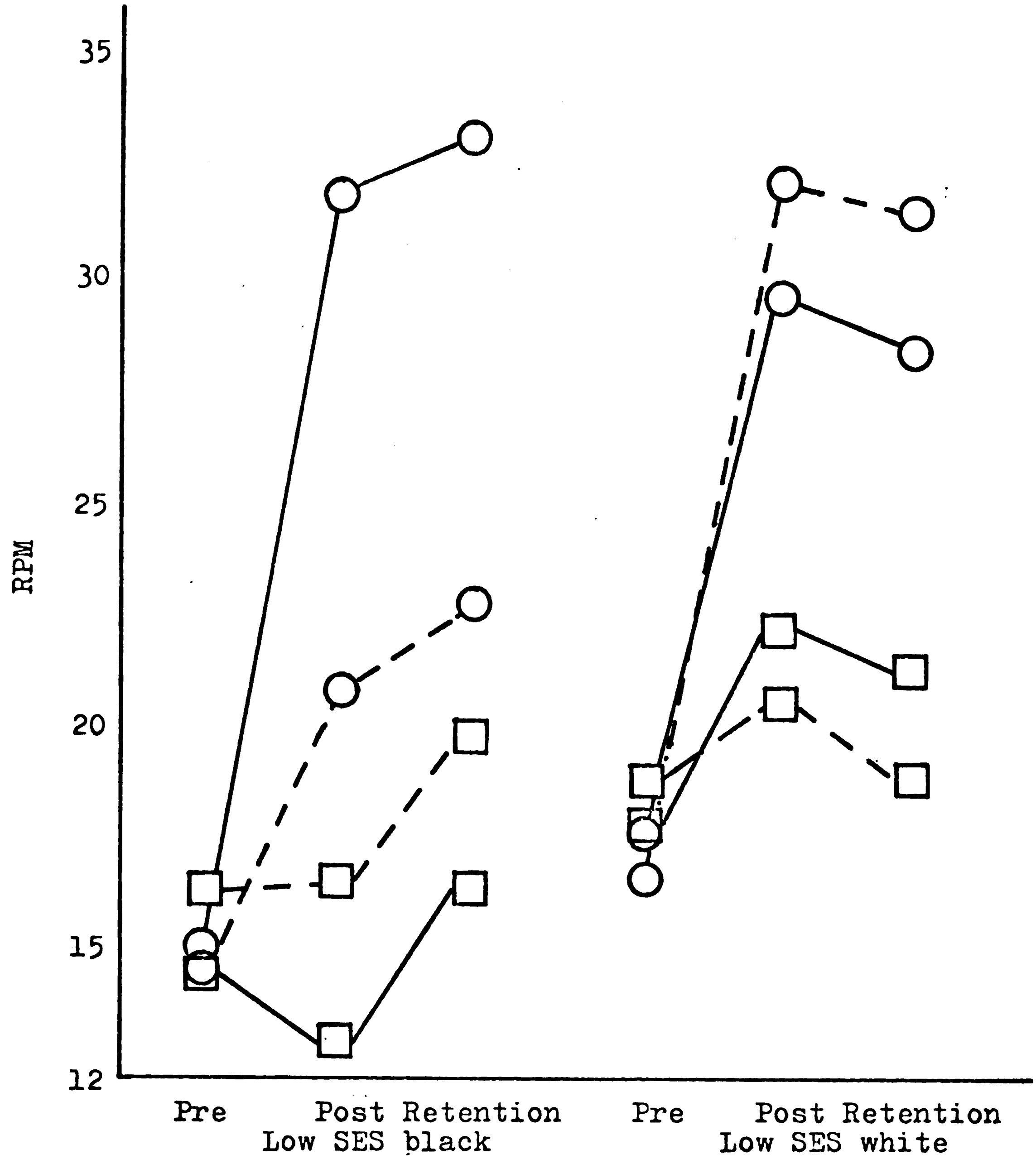


Figure 9. Means of RPM pretest, RPM posttest and RPM retention test.

Experimental-high DS — Control-high DS

Experimental-low DS — Control-low DS

TABLE 5

Analysis of Variance
Using Repeated Measures
on RPM Pretest, RPM Posttest, and RPM Retention Test Scores

Source	df	MS	F
Within subjects			
Trials	2	1472.89	83.20***
Trials X Treatments	2	843.95	47.67***
Trials X DS	2	25.70	1.45
Trials X Race	2	71.56	4.04*
Trials X Treatments X DS	2	30.78	1.74
Trials X Treatments X Race	2	25.12	1.42
Trials X DS X Race	2	43.93	2.48
Trials X Race X Treatments X DS	2	132.43	7.48**
Error (within)	144	17.70	
Between subjects			
Treatments	1	2593.84	36.91***
DS	1	61.00	.87
Race	1	745.54	10.61**
Treatments X Race	1	11.70	.17
Treatments X DS	1	226.20	3.22
DS X Race	1	87.60	1.25
Treatments X DS X Race	1	643.54	9.16**
Error (between)	72	70.28	

Note. All F values above were compared with F values for 1 and 72 df because of Greenhouse and Geisser's correction for within subject sources of variance (Winer, 1962).
*$p<.05$; **$p<.01$; ***$p<.001$

TABLE 6

Post-Hoc Multiple Comparisons of Means for RPM Posttest Scores

	(1)[a] ELW	(2) EHB	(3) EHW	(4) CHW	(5) ELB	(6) CLW	(7) CLB	(8) CHB	Shortest Significant Ranges $p<.01$
Means	32.3	32.0	29.7	22.2	20.8	20.5	16.4	12.9	
(1)		.3	2.6	10.1	11.5	11.8	15.9	19.4	$R_2=6.4$
(2)			2.3	9.8	11.2	11.5	15.6	19.1	$R_3=6.7$
(3)				7.5	8.9	9.2	13.3	16.8	$R_4=6.9$
(4)					1.4	1.7	5.8	9.3	$R_5=7.0$
(5)						.3	4.4	7.9	$R_6=7.1$
(6)							4.1	7.6	$R_7=7.2$
(7)								3.6	$R_8=7.3$
(8)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	
b	———	———	———						
				———	———	———	———		
							———	———	

TABLE 7

Post-Hoc Multiple Comparisons of Means for RPM Retention Test Scores

	(2)[a] EHB	(1) ELW	(3) EHW	(5) ELB	(4) CHW	(7) CLB	(6) CLW	(8) CHB	Shortest Significant Ranges $p<.01$
Means	33.4	31.7	28.5	22.7	21.3	19.8	18.8	16.3	
(2)		.7	4.9	10.7	12.1	13.6	14.6	17.1	$R_2=9.4$
(1)			3.2	9.0	10.4	11.9	12.9	12.2	$R_3=9.8$
(3)				5.8	7.2	8.7	9.7	12.2	$R_4=10.1$
(5)					1.4	2.9	3.9	6.4	$R_5=10.2$
(4)						1.5	2.5	5.0	$R_6=10.4$
(7)							1.0	3.5	$R_7=10.5$
(6)								2.5	$R_8=10.6$
(8)									
	(2)	(1)	(3)	(5)	(4)	(7)	(6)	(8)	
b	———	———	———						
		———	———	———					
			———	———	———	———	———		
				———	———	———	———	———	

[a] E-Experimental; C-Control; H-High Digit Span; L-Low Digit Span; B-Black; W-White.

[b] Any two treatment means not underscored by the same line are significantly different at the .01 level. Any two treatment means underscored by the same line are not significantly different at the .01 level.

results significantly ($p<.01$) different from their control groups. The high DS white group was 7.2 points higher than its control group, which was not significant at the .01 level. The difference was significant at the .1 level however. The low DS black group was only 2.9 points above its control group and the difference was not significant.

The results indicate that the experimental treatment increased the mean scores for the high and low DS white groups and for the high DS black group when compared with their respective control groups. It did not appreciably increase the score of the low DS black group over its control group.

Hypothesis 4. The means of the high DS groups are significantly higher than the means of the low DS groups.

Table 5 reports that there was no over-all effect in the repeated measures analysis attributed to DS level. The analysis of variance for RPM posttest (Table 8) and RPM retention test (Table 9) also reported no over-all significant differences attributed to DS level. However, for all three analyses, there was a significant interaction for the Race X DS X Treatment source of variation: $p<.01$ for repeated measures, $p<.001$ for RPM posttest analysis, $p<.01$ for RPM retention test analysis. This interaction for the RPM posttest data is shown in Figure 10. The figure shows the interaction of the DS level with the experimental

TABLE 8

2 X 2 X 2 Analysis of Variance
for RPM Posttest Scores

Source	df	MS	F
Race	1	621.6	21.2***
DS	1	59.5	2.0
Treatment	1	2257.8	76.9***
Race X Treatment	1	30.0	1.0
Race X DS	1	94.6	3.2
Treatment X DS	1	132.6	4.5*
Race X Treatment X DS	1	435.6	14.8***
Error	72	29.4	

TABLE 9

2 X 2 X 2 Analysis of Variance
for RPM Retention Test Scores

Source	df	MS	F
Race	1	82.0	1.3
DS	1	52.8	1.2
Treatment	1	2010.0	31.4***
Race X Treatment	1	0.0	0.0
Race X DS	1	78.0	1.2
Treatment X DS	1	90.3	1.5
Race X Treatment X DS	1	495.0	7.8**
Error	72	63.6	

*$p<.05$
**$p<.01$
***$p<.001$

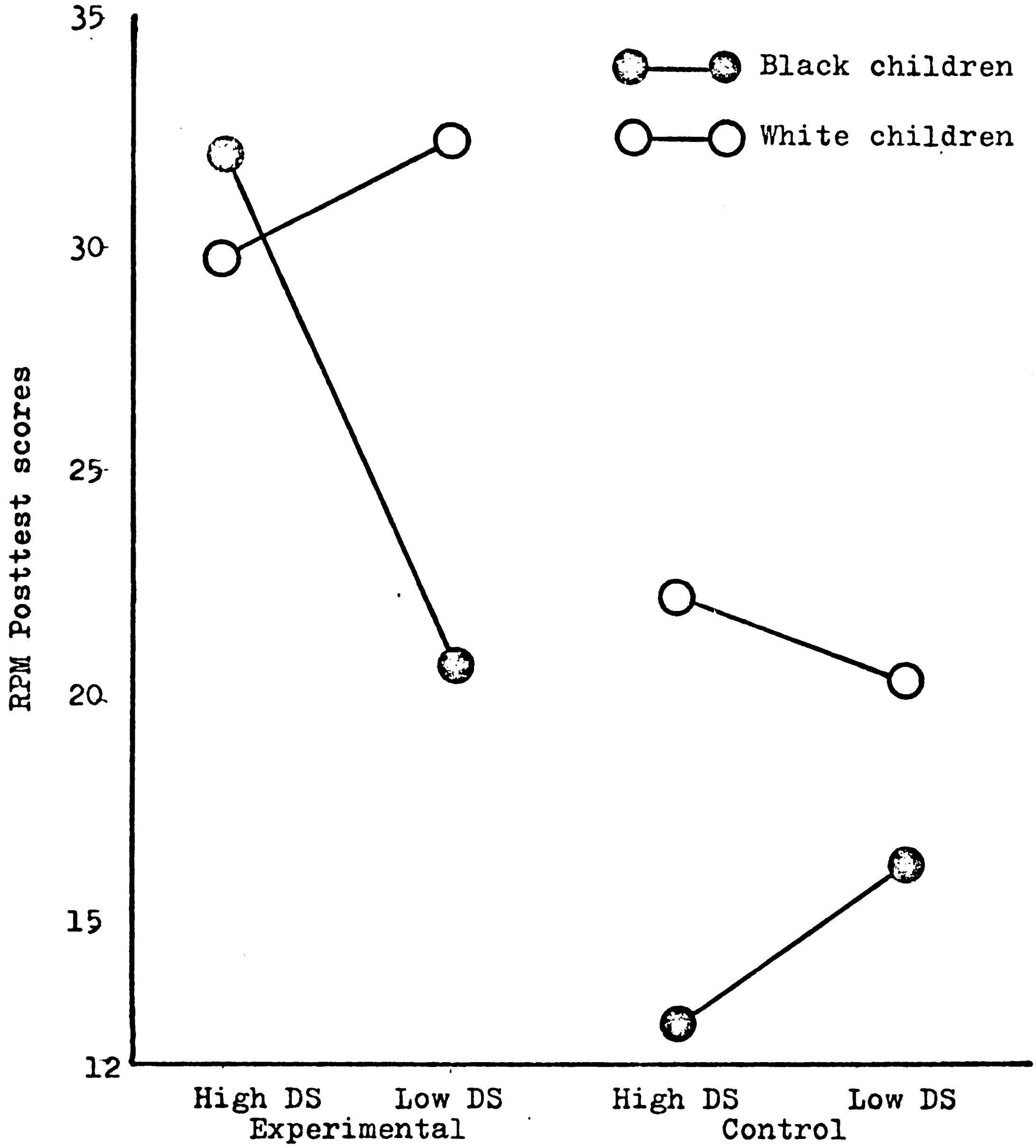


Figure 10. Race X Treatment X Digit Span Interaction using RPM posttest scores.

treatment for the two different races. There was no significant difference between the high and low DS experimental group for the white children. Their DS scores had no effect on how well they learned to solve the RPM. But this was not the case with the black children. The high DS black experimental group had significantly higher means than the low DS black experimental group.

Hypothesis 5. There is no significant difference between the means of the low SES white and the low SES black groups on the RPM.

Table 10 reports the analysis of variance for the RPM pretest. The only significant factor was Race ($p < .001$). The over-all means of the white and black samples before the experimental treatment were 17.75 and 14.85 respectively. Thus the white sample was 2.9 points higher before the experimental treatment began. The posttest analysis of variance reports Race still significant ($p < .001$; see Table 8). The over-all RPM posttest mean was 26.2 for all the white children and 20.5 for all the black children. But by the RPM retention test, no significant difference for Race was found (see Table 9). The mean was 25.1 for all the white children and 23.1 for all the black children. The reason for the 1.1 drop for the white children and the 2.6 point gain for the black children between the RPM posttest and RPM retention test is discussed in the next chapter.

TABLE 10

2 X 2 X 2 Analysis of Variance
for RPM Pretest Scores

Source	df	MS	F
Race	1	168.2	14.8***
DS	1	1.8	.2
Treatment	1	20.0	1.8
Race X Treatment	1	1.8	.2
Race X DS	1	5.0	.4
Treatment X DS	1	24.2	2.1
Race X Treatment X DS	1	.2	0.0
Error	72	11.4	

***$p < .001$

The differences in scores between the races were not an over-all effect: The locus of the differences was in the low DS black group. The mean for each group on the RPM posttest and RPM retention test indicate that there was no significant difference between the three highest scoring groups (see Tables 6 and 7), which were the low DS white experimental, the high DS white experimental, and the high DS black experimental. The low DS black experimental children did not make similar gains. This lack of improvement partially explains the lower over-all scores for the black children.

Hypothesis 6. There is no significant difference between the means of the pure control and Hawthorne

TABLE 11

Hawthorne Versus Pure Control
Analysis of Variance
for RPM Posttest Scores

Source	df	MS	F
Treatment: Hawthorne vs. Pure	1	126.0	4.1
Race	1	442.3	14.4***
DS	1	9.1	.3
Race X Treatment	1	2.1	.1
Race X DS	1	70.3	2.3
Treatment X DS	1	46.3	1.5
Race X Treatment X DS	1	1.1	.04
Error	32	31.6	

TABLE 12

Hawthorne Versus Pure Control
Analysis of Variance
for RPM Retention Test Scores

Source	df	MS	F
Treatment: Hawthorne vs. Pure	1	115.6	1.5
Race	1	39.0	.5
DS	1	2.5	0.0
Race X Treatment	1	153.1	2.0
Race X DS	1	91.0	1.2
Treatment X DS	1	90.0	1.2
Race X Treatment X DS	1	7.3	.1
Error	32	76.7	

***$p < .001$

control groups on the RPM.

The analyses of variance for RPM posttest and RPM retention test report no significant difference between the pure control and Hawthorne control groups (see Tables 11 and 12). In both tests the pure control group did slightly better than the Hawthorne control group, although the difference was not significant.

Sex Differences

Table 13 reports the means and standard deviations of the RPM pretest, RPM posttest, and RPM retention test by sex. There were no significant sex differences found in the study.

TABLE 13

Means and Standard Deviations of RPM Data by Sex

		Male N=13	Female N=7
White experimental	pretest	17.85 (3.39)	15.71 (3.28)
	posttest	30.92 (5.25)	31.14 (4.12)
	retention test	30.08 (5.15)	30.14 (8.27)
White control		Male N=13	Female N=7
	pretest	19.00 (3.58)	17.29 (1.48)
	posttest	21.46 (5.83)	21.14 (4.19)
	retention test	18.92 (8.29)	22.14 (8.13)
		Male N=13	Female N=7
Black experimental	pretest	14.15 (3.46)	15.14 (3.04)
	posttest	26.38 (8.59)	26.43 (6.16)
	retention test	28.62 (9.58)	27.00 (7.05)
Black control		Male N=9	Female N=11
	pretest	15.22 (2.39)	15.18 (3.46)
	posttest	13.56 (4.22)	15.55 (6.75)
	retention test	17.78 (8.47)	18.27 (9.08)

Note. Mean scores are followed by the standard deviation in parentheses.

Intercorrelations Among the Variables

Table 14 and Table 15 report the intercorrelations among DS, RPM pretest, RPM posttest, and RPM retention test. The experimental groups are above the diagonals; the control groups are below the diagonals. The coefficients in the intercorrelation matrices in Table 14 are based on a sample size of 10. Because the sample size was so small, the sampling error was large. For example the correlation between DS and RPM retention test (t_3) for the high DS white experimental group in Table 14 was .30 (N=10). The correlation for the comparable low DS white experimental group was .78 (N=10). When the high and low DS white groups were combined in Table 15, the correlation between DS and RPM retention test falls to -.02 (N=20). What appeared to be a significant correlation between DS and RPM retention test for the high and low DS white children vanishes when the groups were combined. Because of the instability of the intercorrelation matrices in Table 14, only the matrices in Table 15 with high and low DS children combined were examined.

The correlation between DS and RPM posttest and the correlation between DS and RPM retention test differed for the two races. For the low SES white sample the correlations were .02 between DS and RPM posttest and -.02 between DS and RPM retention test. For the low SES black the correlations were .67 between DS and RPM posttest

TABLE 14

Intercorrelations Between DS,
RPM Pretest, RPM Posttest, RPM Retention Test Scores

		DS	t_1	t_2	t_3
Black-high DS					
	DS		.16	.02	.17
RPM Pretest	(t_1)	.27		.00	.04
RPM Posttest	(t_2)	.41	.92		.90
RPM Retention test	(t_3)	.43	.86	.94	
Black-low DS		DS	t_1	t_2	t_3
	DS		.31	.51	.63
	t_1	-.08		.30	.34
	t_2	-.26	.83		.83
	t_3	.20	.38	.57	
White-high DS		DS	t_1	t_2	t_3
	DS		.81	.67	.30
	t_1	-.23		.49	.36
	t_2	-.35	.87		.04
	t_3	-.41	.70	.88	
White-low DS		DS	t_1	t_2	t_3
	DS		.41	.45	.78
	t_1	-.53		.29	.58
	t_2	-.52	.63		.48
	t_3	-.02	.32	.74	

Note. Experimental groups are above the diagonals: control groups are below the diagonals. Each coefficient is based on an N of 10. Because sample size is so small, sampling error is large; e.g., if true correlation is .60, the 95% confidence intervals are -.05 to .89 (Ebel, 1965).

TABLE 15

Intercorrelations Between DS,
RPM Pretest, RPM Posttest, and RPM Retention Test Scores

		DS	t_1	t_2	t_3
Black N=20					
	DS		.14	.67	.25
RPM pretest	(t_1)	-.20		.19	.25
RPM posttest	(t_2)	-.18	.88		.94
RPM retention test	(t_3)	.01	.63	.73	
White N=20		DS	t_1	t_2	t_3
	DS		.46	.02	-.02
	t_1	-.59		.32	.41
	t_2	-.12	.73		.35
	t_3	-.02	.46	.80	

Note. Experimental groups are above the diagonals: control groups are below the diagonals. Each coefficient is based on an N of 20.

and .63 between DS and RPM retention test. These coefficients agree with the interactions reported earlier between Race, Treatment, and DS. For the white children the DS level did not influence the results of RPM posttest or RPM retention test. Both the high and low DS groups learned equally well. For the black groups, the high DS group made larger gains than the low DS group. The higher coefficients for the black group in Table 15 between DS and RPM posttest and between DS and RPM retention test are indicative of this interaction between Race, Treatment, and DS.

Reliability of the Raven's Progressive Matrices

The reliability of the RPM can be examined in the coefficients below the diagonals (the control groups) in Table 15. Because these groups received no experimental treatment, the intercorrelations between the RPM pretest (t_1), RPM posttest (t_2), and RPM retention test (t_3) give a reliability measure of the RPM. The reliability coefficients for the low SES black control group (N=20) were .88, .63, and .73. The reliability coefficients for the low SES white control group (N=20) were .73, .46, and .80. The coefficients were high enough to indicate that the test was reliable.

Results of Experimental Treatment for Middle SES Whites

Table 16 reports the means and standard deviations for DS, RPM pretest, and RPM posttest for the middle SES white children. Originally the middle SES white group was to be included in the analysis with the low SES black and low SES white; however, children scoring low on RPM with at least one college educated parent were difficult to find. Only ten children scored low in a middle SES white sample of approximately 150 children. Because there were so few children, no control group was used, and the children were not divided into high and low DS groups. A retention test was given to only two of the ten children in this group because it was too late in the school year to test the other children.

The results reported in Table 16 indicate that the middle SES white children had RPM posttest scores that were similar to the high DS black experimental children and the high and low DS white experimental children.

TABLE 16

Data for Middle SES White

Student number	DS score	RPM pretest	RPM posttest	RPM retention test
1	35	22	30	
2	26	19	30	
3	60	21	35	
4	40	21	36	42
5	36	11	30	
6	40	23	36	
7	35	18	22	
8	40	13	39	
9	30	23	40	
10	30	18	30	37
Mean	37.2	18.9	32.8	
Standard deviation	(12.1)	(6.1)	(10.5)	

Intercorrelation Matrix

		DS	t_1	t_2
	DS		.08	.26
RPM pretest	(t_1)			.21
RPM posttest	(t_2)			

Note. Children scoring low on the RPM with at least one college educated parent were difficult to find. These ten students were from four diffferent schools and an original sample of approximately 150 children with one or more college educated parent. There is no control group for the middle SES white. The retention test was given to only two students becasue it was too late in the year to test the other students.

Summary

The first two hypotheses dealt with the form of the scatter-diagram between Basic Learning Ability and Intelligence. The pattern which Jensen believed to exist between the two concepts was not found. According to Jensen, there should be no scores falling in the low Basic Learning Ability-high Intelligence quadrant. In this study in each SES sample, significant numbers of these scores were found in this particular quadrant. The correlations between the Digit Span test and Raven's Progressive Matrices were: low SES white, .10; low SES black, .22; middle SES white, .34. Jensen hypothesized that the correlations would be below .20 for the low SES groups and between .50 and .70 for the middle SES group. While the middle SES correlation was not as high as .50, the correlation for the middle SES white sample was higher than the correlations for the two low SES samples.

The experimental treatment interacted differently with the high and low digit span experimental groups for the two races. In the low SES white sample, both the high and low digit span groups had scores on the Raven's Progressive Matrices' posttest significantly greater than their respective control group. In the low SES black group, only the high digit span experimental group had Raven's Progressive Matrices' posttest scores larger than its control group. The low digit span group did not gain from the experimental treatment. Results were similar for the Raven's *Progressive Matrices'* retention test, which was given one month after the posttest.

Ten middle SES white children also received the experimental treatment. Enough children for a control group in the middle SES white sample could not be found because there were not enough children who scored low on the Raven's *Progressive Matrices* and had at least one college educated parent. Middle SES white children made gains similar to the low and high digit span low SES white sample and the high digit span low SES black sample.

CHAPTER V

DISCUSSION

This study had two goals: First, to replicate some descriptive findings reported by Jensen, and second, to test the degree that these descriptive findings can be modified by an experimental treatment. Jensen believes that different SES levels and different races have different patterns of abilities. He feels that the major cause for these differences resides not in the environment, but in genetic differences between different populations. This study holds that the differences in ability between SES levels are the results of environmental differences. Unfortunately, research does not always serve to give answers, but often serves only to complicate the old answers and to raise new questions.

One point that has been clearly made by this study is that scores on the Raven's _Progressive Matrices_ can be substantially increased as the result of a training program. Jensen states in his _Harvard Educational Review_ article (1969, p. 101), "I have found no studies that demonstrated gains in relatively noncultural or nonverbal tests like Cattell's _Cultural Fair Tests_ and Raven's

Progressive Matrices." This study clearly finds these gains for the middle SES white children, the low SES white children and the high digit span low SES black children. The posttest scores for these groups were all within four points of the middle SES white pretest scores. The improvements were not only evident in the posttest but were maintained in the retention test given one month later.

The reason that this study succeeded where other attempts have failed was because the training was developed using task analysis to look at how good RPM solvers actually operate. The children were trained not directly on the RPM tasks, but on the strategies and concepts prerequisite to the RPM solution. The development of the training program was not based on speculation, but on observation of the strategies and concepts used by successful and unsuccessful RPM solvers. The success of this study demonstrates that there are no short cuts to learning: The prerequisite knowledge and skills must be taught if success is to be achieved.

Correlations between Digit Span and Progressive Matrices

Jensen reports correlations between Intelligence (IQ) and Basic Learning Ability (BLA) for different SES levels. He reports that the correlation between BLA (in this study measured by a digit span test) and IQ (in this study measured by Raven's Progressive Matrices) differs for the middle and low SES subjects. He claims that for the middle SES the correlation is between .50 and .70; for the low SES level he finds correlations between 0 and .20.

The results of this study do not find the differences between the SES levels to be this large. The correlations were: .23 for the low SES black children (N=105); .10 for the low SES white children (N=84); .34 for the middle SES white children (N=79). However when the correlations between the Digit Span (DS) test and the Raven's Progressive Matrices (RPM) for each classroom are examined, the values are very inconsistent. In the low black group, three of the four classrooms had correlations of essentially zero: The fourth had a correlation of .33. In the low white group, two of the three classrooms were essentially zero and the third was .20.

The middle SES white children were from two different schools with two classrooms in each school. One school was the University of Florida laboratory school. This school contains the most priviledged group of children in the middle SES white sample because most of the children's fathers have attended graduate or professional schools. In each classroom at the laboratory school the correlations between DS and RPM were .22. The over-all correlation when the classrooms were combined was also .22 (N=46). The other middle SES white school was in a suburban area. At this school the parents had college educations, but not advanced degrees. The correlations between DS and RPM for each classroom in this school were .44. The over-all correlation for this school was .44 (N=33). The reasons for the differences in correlations between these two schools is unclear.

Testing done last year in Haslett, Michigan, showed similar correlations. The correlation was .30 for second graders (N=70), .39 for fourth graders (N=48), and .40 for sixth graders (N=45). The majority of these children were from the middle SES level, although many of their parents were probably not college educated.

The correlations for the low SES levels were generally lower than the correlations for the middle SES level, although the middle SES correlations were not as high as those that Jensen found. It is surprising that the group of children with the highest SES level, the University of Florida laboratory school children, had a correlation of only .22. This correlation was almost identical to the correlation of the low SES black children (.23).

The Patterns of the Scatter-diagram

Jensen hypothesizes that although a high BLA, as measured by DS, is a necessary prerequisite for high IQ, it is not sufficient for high IQ. Jensen believes that because IQ is genetically independent of BLA, it is possible for an individual to have a high BLA and yet have a low IQ. In this case, the low IQ score can be attributed to the genes associated with IQ. However, based on his theory, an individual possessing both low BLA and high IQ cannot exist because high IQ is functionally dependent upon high BLA. Figures 6, 7, and 8 (pp. 52, 53, 54) show that in each SES level there were children whose scores placed them in the low DS-high RPM quadrant. There was no evidence in any SES level that high DS was necessary for high RPM or that an individual with a low DS could not have a high RPM score. Nor did there seem to be any absolute lower limit of DS scores, below which the RPM score could not be above.

Genetic Independence of IQ and BLA

Jensen believes that IQ and BLA spring from genetically separate sources. This study gives no evidence to support or refute this claim. The phenotypic expression of these two constructs (DS and RPM) are correlated. Jensen's explanation for this correlation is that BLA is necessary for IQ. This is not a causal relationship, which would imply not only necessity but also sufficiency. This study shows that low BLA does not necessarily make high IQ impossible. A better interpretation than the necessary relationship between IQ and BLA is the common variance explanation: BLA and IQ are correlated because the DS and RPM share a common variance with general intelligence.

Experimental Treatment

A small group was selected from each of the three populations to participate in an experimental treatment designed to raise their RPM scores. These children originally had scored low on the RPM, which usually meant a score of less than twenty. An equal number of children with high and low DS scores received the treatment, which lasted three and one-half hours over seven different one-half hour sessions. After the treatment, the experimental children and an equivalent group of control children were given a RPM posttest. A month after the RPM posttest, a RPM retention test was given to all the children that took the RPM posttest.

Differences in the RPM Pretest

As discussed previously, there were large differences among the RPM pretest scores of the three populations in the study. In order that comparisons could be made across groups, the mean scores for each of the populations were kept as equivalent as possible. But by keeping the absolute values equivalent across groups, the relative position within each group differed for each of the three populations. In the low SES black group, the mean of the experimental group was 14.5 (N=20) which is .1 standard deviation above the over-all group mean of 13.9 (N=105). For the low SES white group, the mean of the experimental group was

17.0 (N=20), which is .9 standard deviation below their group mean of 24.6 (N=84). For the middle SES white group, the mean of the experimental group (N=10) was 18.9, which is 2.0 standard deviations below their group mean of 34.1 (N=79). Thus even though the absolute scores for each of the three populations were similar, their relative position within their populations differed.

Posttest Results

The results of the experimental treatment were shown in Figure 9 (p. 57). Both the low and high DS groups of the low SES white experimental children improved. The comparable control groups did not. The middle SES white experimental children also improved their scores on the posttest. But the pattern of scores was different for the low SES black children. The high DS group had an increase in scores similar to the white students. In fact by the retention test this group had a higher mean score than either of the low SES white experimental groups. But the low DS black group did not improve as much as the other experimental groups. Most of the gain can be attributed to two students who had a pattern of scores similar to the high DS black students (Students 15 and 16; see Appendix A). Both of these children were retested on the DS to see if they had been misclassified. One score changed from 26 to 36 and the other changed from 40 to 51. Thus the former student was definitely classified correctly, while the latter might have been misclassified. However regression effects and increased rapport with the teacher might have caused the increases. It is the author's belief that these children were not misclassified, and while their RPM posttest scores were similar to those of the children in the high DS group, they belonged in the low DS group. When these two children are removed from the analysis, the mean of the posttest scores drops from 20.8 to 18. Thus after three and one-half hours of coaching, the average gain was less than four points for the eight remaining low DS black children. This is discouraging when compared with gains of 17 points for the high DS black children.

Correlations between DS and RPM Posttest

The effect of the DS level on ability to learn concepts tested by the RPM can be seen in the correlation between DS and RPM posttest. The correlation between DS and RPM pretest should be small because the range of the RPM scores for the children selected for the experimental treatment was small. The correlation between the DS and the RPM pretest was -.05 for the low SES white children (N=40), and -.02 for the low SES black children (N=40). If the correlation were positive for the experimental groups after the experimental treatment, the indication would be that DS is a correlate of learning for the RPM. The differences between the low SES white and black experimental groups can be seen in Table 15 (p. 69). The correlation was .67 for the black children (N=20), and .02 for the white children (N=20). These findings were consistent with the change in the mean scores between the RPM pretest and RPM posttest. The RPM posttest scores for the high DS low SES black children were higher than the RPM posttest scores for the low DS low SES black children, while there was no significant difference between the RPM posttest means of the high and low DS low SES white children. It should be noted that the correlations in Table 14 (p. 68) are unstable due to the restricted range of the DS scores and to sampling error. A correlation based on the largest N and the greatest range of variance (Table 15) gives a more accurate picture of the relationship between the variables.

Rapport and Personality Differences

In general, the children seemed to enjoy the training sessions. It was a change in their daily routine and they were given special attention.[3] Of the thirty black children, twenty experimental and ten Hawthorne control,

---

[3]To eliminate the possibility that gains could be attributed to a Hawthorne effect, the trainers worked with an equivalent group of children on verbal skills unrelated to the RPM. There were no significant differences between the pure control groups and the Hawthorne control groups.

there were only three problem children (Students 21, 29, and 16; see Appendix A); of the thirty low SES white children there were three problem children (Students 12, 20, and 26); of the ten middle SES white there were three problem children (Students 2, 3, and 9). The trainers defined problem children as those whose attitudes and behaviors hindered their coaching sessions.

There were obvious personality differences between the groups. The trainers found the black children the most interesting children to work with. They were a very heterogeneous group, some happy, some sad, some talkative, some quiet. As a group, they were more spontaneous, affectionate, and tried to please the trainers more than the low SES white children. The low SES white children were more homogeneous, very controlled, and more peer-oriented. The trainers found these children less satisfying to work with than the black children. It is interesting to note that the behavior of the black children would cause difficulty in a classroom, as such spontaneous and affectionate children would be more difficult to handle in groups of thirty. The low SES white children on the other hand would be easier to work with in large groups. No generalizations can be made about the middle SES white children. Some were difficult to work with because they were so independent. Others lacked confidence in their ability to learn.

These personality differences seem to explain the small changes from the RPM posttest to the RPM retention test. Figure 9 (p. 57) shows that all four black groups increased their mean scores on the RPM retention test: The mean gain was 2.6 points. All four low SES white groups had smaller means on the RPM retention test: The mean drop was 1.1 points. For both the RPM posttest and the RPM retention test, the testing conditions were not ideal. There was some commotion in the room which differed somewhat for each group. The black children argued over who was going to sit next to the trainer and demanded individualized attention from her. If they did not receive it quickly, they acted out or came right to her. The increase in scores for the black children was interpreted as the result of the children's attempt to please the trainer.

The disorder in the low SES white group was less teacher-oriented. The children scrambled for seats also, but usually took one near a friend rather than near the trainer. Several of the students raced with each other to finish the test. This type of behavior explains the small drop in scores for the low SES white. For both races, however, the differences between the RPM posttests and the RPM retention tests were not significant.

Impressions of Low Versus High DS Children

Even though the trainers did not know which children were in the high DS group and which were in the low DS group, they could readily pick the high DS students in the black group. In the white group they could not do this. The trainer who did most of the work with the children was presented with two names from the same sample, one with a high DS and one with a low DS. The trainer was asked to pick the better student in the pair. There were fifteen pairs of children in both the low SES black and low SES white groups composed of the ten children in the Hawthorne control groups and the twenty experimental children. For the black children, the trainer said 11 of the high DS children were better students, 1 of the low DS students was better, and was undecided for three pairs. For the white children, 7 of the high DS children were better students, 6 of the low DS children were better, and for 2 the trainer was undecided. The low DS measure had obvious behavioral correlates for the black children, which were not present for the white children.

The Meaning of Low DS for the Black Children

First, a low DS does not doom a black child to failure. Child 16 (see Appendix A) had a low DS and improved his score. Child 38, in the low DS pure control group, changed her score from 19 on the RPM pretest, to

27 on the RPM posttest, to 43 on the RPM retention test. So there were exceptions. But generally, the low DS black children did not do well by any measure of success.

The low DS black children generally were very quiet children. One child's teacher was surprised that he spoke to the trainers at all (Student 12), since she claimed that the child had not spoken much since the beginning of the school year. Therefore, it was difficult for the coaches to know if the children really understood the material. Because there was doubt whether the child understood the material, many times the coaches would go over the materials so much that the child would reach the criterion through what appeared to be memorization alone.

There are two explanations as to why these low DS black children did not learn. Jensen's explanation would be that these were children that genetically did not have the ability to learn the concepts involved in the RPM. They preferred to memorize everything because this is their stronger ability. Because the training session demanded an abstract ability which the children did not possess, they failed. The failure of the children to learn these abstract skills from the experimental treat ment is further evidence that, for these low SES black children, abstract intellectual skills are unaffected by intervention.

The second explanation states that the children did not learn because the three and one-half hours training was not sufficient time to break old learning habits and instill new ones. Eight years of deprivation could not be overcome in three and one-half hours. The low DS children used memorizing techniques in attacking the problems not because they were better at memorizing, but because they knew no other way of attacking the problems. It should be pointed out that based on these children's DS scores, these children were not very good at BLA. The high DS black children did not resort to memorizing techniques even though their BLA was better than the low DS black children. The high DS black attacked the problems in the same way that the low SES and middle SES white children attacked the problems.

It is the author's belief that in this experiment BLA acts as a correlate of IQ. It is not a perfect correlate of IQ (.52 between DS and the WISC [Wechsler, 1949]) but is an indication of intellectual level. All the white children learned because these children were not as deprived as the black children. They had learned enough skills from their environment to take advantage of the coaching sessions and learn. For the black children the situation was different. The high DS children were able to overcome any deficiencies in their educational background because they were more intelligent children.

The low DS children were not able to improve because they were not as intelligent. In other words, the training was useful only if certain prerequisite skills were present: If these skills were not present, the training could not be successful.

The children needed skills in at least two areas in order to profit from the training procedure. First, they had to understand what they were supposed to do. The low DS low SES black children seemed quite content to guess at the solution to a problem probably because they did not understand the problem. Thus part of the entering behavior needed would be some experience with the idea of a problem and that problems have solutions. Another skill lacking in the low DS low SES black group was the understanding of a concept. The idea that knowledge gained on one problem could be applied to the next problem seemed to be foreign to these children. This is perhaps why many of them attempted to memorize the solutions to a particular problem. Both of these skills are basic prerequisite skills that the trainers assumed the children possessed when they started the training procedure. Perhaps these basic skills were not present with the low DS low SES black children, causing the training to be less than successful.

Creating a new theory to explain the data is difficult. The two populations, low SES white and low SES black reacted differently to the treatment. Figure 11 is a theoretical attempt to explain the data. The horizontal axis goes from deprived on the left to enriched on the right. The low SES black children are considered more intellectually deprived than the low SES white children. Here the crucial variable is not race, but intellectual deprivation. The vertical axis extends from low to high IQ. DS acts only as a correlate of IQ and not as a separate ability. Thus a low DS indicates a low IQ. Any individual above the curved line can learn in the training session. Any individual below and to the left of the curved line will not learn. Thus success depends on a combination of both the environment and IQ.

Comments on Jensen's Harvard Educational Review Article

In his article in the Harvard Educational Review, Jensen discusses, the genetic and environmental influences of IQ. The following comments are based partially on his study.

Difficulty in Developing Comparable Scatter-diagrams

Figure 1 (p. 11) shows the hypothetical different scatter-diagrams of BLA and IQ in low SES and upper-middle SES groups (Jensen, 1969a). There are several problems with diagrams that make their interpretation difficult.

It is assumed the $\bar{x}$'s refer to the mean of each SES

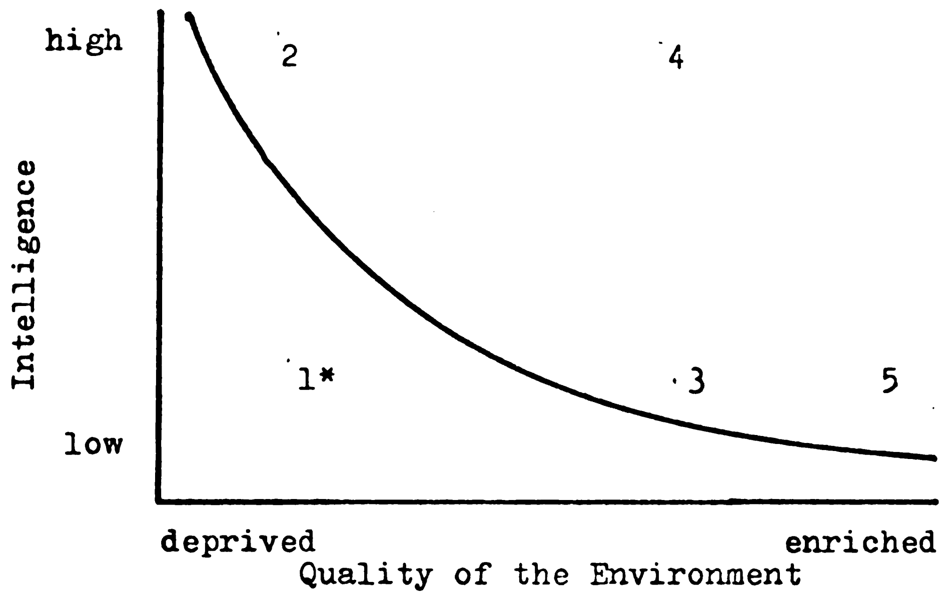


Figure 11. Post-hoc explanation of the data. The horizontal axis goes from deprived on the left to enriched on the right. The low SES blacks are considered more intellectually deprived than the low SES whites. Here the crucial variable is not race, but intellectual deprivation. The vertical axis extends from low to high IQ. DS acts only as a correlate of IQ and not as a separate ability. Thus a low DS means a low IQ. Any individual above the curved line can learn in the training session. Any individual below and to the left of the curved line will not learn. Thus success depends on a combination of both the environment and intelligence.

*1. Low SES black-low DS
2. Low SES black-high DS
3. Low SES white-low DS
4. Low SES white-high DS
5. High SES white-low DS

group, and not to the over-all mean of both SES groups combined. The density of the low SES scatter-diagram gives three times as many scores above the mean for BLA as there are below the mean; it also gives three times as many scores below the mean for IQ as there are above the mean. If this is actually the case the distribution for the low SES scatter-diagram must be extremely skewed toward the high IQ scores and extremely skewed toward the low BLA scores. This skewed distribution for IQ and BLA for low SES blacks has not been found. Kennedy, Van De Riet, and White (1963) found that the distribution of the Stanford-Binet scores for low SES black children gave a normal curve with a mean of 80.7 and a standard deviation of 12.4 as opposed to data upon which the test was standardized, which gave a mean of 101.8 and a standard deviation of 16.4. The normal curve for the low SES black children has no skew in it, but was a normal curve with a smaller mean and a smaller standard deviation.[4]

This creates a problem in comparing different scatter-diagrams when the means of the groups are so different. For example on the RPM pretest scores, the low SES black mean is 2.2 standard deviations below the low SES white mean and 4 standard deviations below the mean of

---

[4]It should be noted that Jensen makes comparisons between SES levels, not races. The distribution for other low SES populations may differ from that of the low SES black population.

the middle SES white. High and low RPM pretest scores therefore are not absolutes. Individuals are above or below their group's mean, but often above the mean of the RPM scores for the low SES black are below the mean scores for the low SES white and the middle SES white.

Teaching to Accommodate Different Patterns of Ability

Jensen concludes that because different SES levels have different abilities, each group should be taught by emphasizing tasks that the child can attack using his strong skills. For the disadvantaged child, according to Jensen, this ability is associative learning. Jensen is not very specific about how this should be done, and in a 123 page article he devotes only one and one-half pages to the possibility of a new type of curriculum (1969a). When Jensen uses the word "strong" in relation to ability, it must be remembered that this does not mean that the low SES is better in BLA than the middle SES levels. Based on the data for DS scores for the different SES levels, the two SES levels are at best equal (see Table 3; p. 49). When he says better, he means that the low SES group's BLA is better than their abstract intellectual ability. In fact what he is suggesting is that their abstract intellectual ability is so poor that any hope of reaching it is futile, and we should not waste our time trying to reach the children at this abstract level where little can be expected.

Carl Bereiter is even more pessimistic about the possiblity of giving abstract intellectual tools to all children. In his discussion of the Jensen article (1969), he suggests that there are certain intellectual tools certain people can never acquire. For example, let us say that with present teaching techniques, only 10% of the population can grasp a concept in calculus. As we discover new ways to teach this abstract concept, more people will acquire it, but many will never be able to acquire this new skill. Thus Bereiter believes the gap between the "haves" and the "have-nots" will become even greater, as the "haves" improve and the "have-nots" stand still.

What makes Jensen's suggestion for a special curriculum depressing is that it restricts the tools that the low SES can use. As this study has demonstrated, many low SES children can acquire these more abstract conceptual skills. Although it is not in the data, observation of the children demonstrates that the associative intellectual tools that the children brought into the coaching sessions were of no value in solving problems which demanded abstract conceptual tools. How Jensen is to turn associative learning into abstract intellectual skills is not clear. It is naive to believe that it is possible to be an educated individual without abstract conceptual skills.

There are many psychologists who feel that their compensatory education programs have not failed. Robert Spaulding at Duke University (personal communication) has developed a program that emphasized problem solving and discovery learning. The gains for the first year of school are small: The children are one-half year behind the national average. By the end of the second year of school, the children have gained one year but are still one-half year behind the national average. But by the end of the third year, the children have caught up with the national average. This type of program demonstrates not only that compensatory education can succeed, but that success comes only after years of work. In this particular compensatory education program, it took three years to bring the children up to the national average.

Implications for Future Research

The basic question that this study raises is why did the low DS low SES black children fail to gain the concepts necessary to solve the problems on the RPM? Could the children have learned if the coaching procedures had been modified? The coaching procedure could easily have been changed in several ways. First, the total amount of time a child was coached could be increased. Second, the materials could be more concrete. For example actual wooden puzzles could be used instead of the more abstract drawings of the matrices. Third, concrete reinforcement such as candy or money might improve the performance of the children.

Another possible approach might be to give the low DS low SES black children training for the DS test in order to increase their ability to learn. Studies that trained children to do the DS have not been too successful. In a study discussed by Woodworth and Schlosberg (1954), kindergarten children after 78 days of training increased their digit span from 4.4 to 6.4 digits. After a long vacation in which no further practice occurred, they averaged 4.7 digits, almost down to their original level and probably no further above than could be accounted for by growth. So the training showed temporary gains that were not maintained on the retention test. But there is the possibility that training for the DS might improve the children's ability to learn the concepts in the RPM test.

The relationship between DS and the ability to improve on the RPM might become more obvious in the low SES white group if the training procedure reduced the amount of feedback to the student. The lack of difference between the gains made by the low DS and high DS children is probably the result of a uniform improvement for all children which was caused by the thorough training that reached all low SES white children. It is possible that the training procedure held back the high DS low white children because it was too structured. Instead of bringing each child up to a specific criterion, the materials could be presented in a lecture fashion. Each child would not receive as much attention as the present coaching procedure permits and individual differences should be larger. This might increase the correlation between the DS and the gain on the RPM.

By altering the training procedure the curved line in Figure 11 will be changed. Discovering the attributes of different training programs that eliminate or increase individual differences would help clarify learning abilities for various populations. As Cronbach (1967) says "...Our greatest hope for fitting the school to the individual lies in the development of theory that finally marries the differential and experimental approach to learning."

The Meaning of a Genetic Interpretation

It is obvious that the easiest interpretation of the data was a genetic one. The interpretation given in this study was more complex and involved more hindsight and post-hoc analysis. But when the results are not clear-cut, interpretations can be influenced greatly by social values. It is wrong to say that children are genetically different on the basis of a mere three and one-half hours of coaching. There is an analogy between the problem we are faced with and the type I and type II error found in statistics courses. A type I error is committed when a true hypothesis is rejected. A type II error is made when a false hypothesis is accepted. Suppose our hypothesis is that the genetic make-up of different SES levels is so constituted that the middle SES levels are genetically superior to the low SES levels. If this is a true hypothesis and it is rejected, a type I error has been made. But suppose this hypothesis is accepted when it is actually false. This is a type II error of the most serious sort. A population has been labeled genetically inferior when they well may not be. Before making such a mistake, psychologists should continue to prefer the type I error to the type II error until more data are in.

Major Findings of the Study

The study was both successful and unsuccessful in changing scores on the RPM through coaching. It was successful because four of the five experimental groups made significant gains after the experimental treatment. It was unsuccessful because the low DS low SES black group did not improve. Jensen says in his *Harvard Educational Review* article (1969, p. 101), "I have found no studies that demonstrated gains in relatively noncultural or nonverbal tests like Cattell's Culture Fair Tests and Raven's Progressive Matrices." This study clearly finds such gains for the middle SES white children, the low SES white children, and the high DS low black children.

The failure of the low DS low SES black group to improve can be attributed to either environmental deficiency or genetic deficiency: This study can be interpreted either way. What is clear in the study is that reaching these children is a difficult task. Jensen suggests that the tactics of compensatory education should call upon the BLA skills rather than abstract conceptual skills. However it is hard to see how such simple skills could transfer to the abstract skills needed to solve problems such as those in the RPM.

The three hypotheses Jensen developed to explain the pattern of scores for BLA and IQ which he found in the middle and low SES levels (see Figure 1; p. 11) can now be examined in light of the findings of this study. First the pattern of scores Jensen reported was not found. The scatter-diagrams for both the middle and low SES groups were not as different as those hypothesized by Jensen, so perhaps the three hypotheses need to be rethought. Jensen's three hypotheses are stated next along with the findings of the study:

1. The genes determining BLA and IQ assort independently.

This hypothesis was not examined by this study.

2. IQ is functionally dependent on BLA.

No functional relationship was found in the scatter-diagrams between DS and RPM. The two concepts are correlated because each is correlated with IQ.

3. While the genotypic basis for BLA is distributed equally in middle and low SES, the genotypic basis for IQ is lower for low SES than middle SES.

Of course genotypes can never be examined directly. The pretest data show similar distributions for both SES levels for DS and different distributions for RPM. After training for the RPM, four of the five experimental groups had mean scores that were not significantly different from the middle SES white RPM pretest. Only the low DS low SES black children were still below the middle SES white RPM mean. The lack of improvement for these children was interpreted as the result of an interaction between the environment and the intellectual level, not as a result of a genetic deficiency.

Creating a new theory to explain the data is difficult. The two populations, low SES white and low SES black reacted differently to the treatment. Figure 11 (p. 91) is an attempt to explain the data. The horizontal axis goes from deprived on the left to enriched on the right. The low SES black children are considered more intellectually deprived than the low SES white children. Here the crucial variable is not race, but intellectual deprivation. The vertical axis extends from low to high IQ. DS acts only as a correlate of IQ and not as a separate ability. Thus a low DS is one indication of low IQ. Any individual above the curved line can learn in the training session. Any individual below and to the left of the curved line will not learn. Thus success for an intellectual task depends on a combination of both the environment and intelligence.

LIST OF REFERENCES

LIST OF REFERENCES

Anastasi, A. *Differential Psychology.* (3rd ed.) New York: Macmillan, 1964.

Anderson, R. C. Individual differences in problem solving. In R. M. Gagne (Ed.), *Learning and Individual Differences.* Columbus, Ohio: Charles E. Merrill, 1967. Pp. 68-89.

Barratt, E. S. The relationship of the Progressive Matrices (1938) and the Columbia Maturity Scales to the WISC. *Journal of Consulting Psychology*, 1956, *20*, 294-296.

Bereiter, C. The future of individual differences. *Harvard Educational Review*, 1969, *39*, 310-318.

Bereiter, C., & Engleman, S. *Teaching disadvantaged children in the preschool.* Englewood Cliffs, N. J.: Prentice-Hall, 1966.

Burke, H. R. Raven's Progressive Matrices: A review and critical evaluation. *The Journal of Genetic Psychology*, 1958, *93*, 199-228.

Burt, C. The inheritance of mental ability. *American Psychologist*, 1958, *13*, 1-15.

Cattell, J. M. Mental tests and measurements. In W. Dennis, *Reading in the history of psychology.* New York: Appleton-Century-Crofts, 1948. Pp. 347-354.

Cronbach, L. J. *Essentials of psychological testing.* (2nd ed.) New York: Harpar & Row, 1960.

Cronbach, L. J. Heredity, environment, and educational policy. *Harvard Educational Review*, 1969, *39*, 338-347.

Duncanson, J. P. *Intelligence and the ability to learn.* Princeton, N. J.: Educational Testing Service, 1964.

Edwards, A. L. *Experimental design in psychological research.* New York: Holt, Rinehart, and Winston, 1960.

Eells, K., Davis, A., Havighurst, R., Herrick, V., & Tyler, R. Intelligence and cultural differences. Chicago: University of Chicago Press, 1951.

Gagné, R. The acquisition of knowledge. Psychological Review, 1962, 69, 465-568.

Greene, K. B. The influence of specialized training on tests of general intelligence. 27th Yearbook, National Society for the Study of Education, 1928, Part I, pp. 421-428.

Guilford, J. P. The nature of human intelligence. New York: McGraw-Hill, 1967.

Haggard, E. A. Social-status and intelligence: An experimental study of certain cultural determinants of measured intelligence. Genetic Psychological Monographs, 1954, 49, 141-186.

Higgins, C., & Sivers, C. A comparison of Stanford-Binet and Colored Raven Progressive Matrices IQs for children with low socioeconomic status. Journal of Consulting Psychology, 1958, 22, 465-568.

Hodges, W. L., & Spicker, H. H. The effects of preschool experiences on culturally deprived children. In W. W. Hartup & N. L. Smothergill (Eds.), The young child: Reviews of research. Washington, D. C.: National Association for the Education of Young Children, 1967. Pp. 262-289.

Hunt, J. McV. Has compensatory education failed? Has it beem attempted? Harvard Educational Review, 1969, 39, 278-300.

Jacobs, P. I. Programmed progressive matrices. Proceedings of the 74th Annual Convention of the American Psychological Association. New York: American Psychological Association, 1966.

Jacobs, P. I. & Vandeventer, M. Progressive Matrices: An experimental developmental, nonfactorial analysis. Research Bulletin RB-68-3. Princeton: Educational Testing Service, January, 1968.

Jensen, A. R. Learning abilities in Mexican-American and Anglo-American children. California Journal of Educational Research, 1961, 12, 147-159.

Jensen, A. R. Learning ability in retarded, average, and gifted children. Merrill-Palmer Quarterly, 1963, 9, 123-140.

Jensen, A. R. Social class, race, and genetics: Implications for education. American Educational Research Journal, 1968, 5, 1-42. (a)

Jensen, A. R. Intelligence, learning ability, and socioeconomic status. In New approaches to the measurement of intelligence. Symposium presented at the annual convention of the American Educational Research Assocation, Chicago, February 8, 1968. (b)

Jensen, A. R. Patterns of mental ability and socioeconomic status. Paper presented at the annual meeting of the National Academy of Science, Washington, D. C., April 24, 1968. (c)

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

Keir, G. The Progressive Matrices as applied to school children. British Journal of Psychology, Statistical Section, 1949, 2, 140-150.

Kennedy, W. A., Van De Riet, V., & White, J. C., Jr. A normative sample of intelligence and achievement of Negro elementary school children in the Southeastern United States. Monograph of the Society for Research in Child Development, 1963, 28, No. 6.

Ludlow, H. G. Some recent research on the Davis-Eells games. School and Society, 1956, 84, 146-148.

McNemar, Q. Lost: Our intelligence. Why? American Psychologist, 1964, 19, 871-882.

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

Raven, J. C. Guide to using Coloured Progressive Matrices. London: H. K. Lewis, 1956.

Raven, J. C. Guide to the Standard Progressive Matrices. New York: The Psychological Corporation, 1960.

Rohwer, W. D., Jr. Social class differences in the role of linguistic structures in paired-associate learning. Cooperative Research Project No. 5-0605, U. S. Office of Education (1967).

Stake, R. Learning parameters, aptitudes, and achievements. Psychometric Monographs, 1961, No. 9.

Thurstone, L. L. Primary mental abilities. Psychometric Monographs, 1938, No. 1.

Tyler, L. A. The psychology of human differences. (3rd ed.) New York: Appleton-Century-Crofts, 1965.

U. S. Commission on Civil Rights. Racial isolation in the public schools. Vol. 1. Washington, D. C.: U. S. Government Printing Office, 1967.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

Wiseman, S. Symposium on the effects of coaching and practice in intelligence tests. British Journal of Educational Psychology, 1954, 24, 5.

Woodrow, H. The ability to learn. Psychological Review, 1946, 53, 147-158.

Woodworth, R. S. and Schlosberg, H. Experimental psychology. New York: Henry Holt and Company, 1954.

APPENDIX A

Data for Experimental and Control Children

Student number	Sex	DS score	RPM pretest	RPM posttest	RPM retention test
White - high DS - experimental					
1	M	95	25	39	33
2	M	58	20	25	35
3	F	74	19	31	36
4	M	63	17	26	32
5	F	68	20	29	26
6	M	70	17	28	20
7	F	57	13	34	28
8	F	50	15	26	31
9	M	57	19	33	21
10	M	50	12	26	23
White - low DS - experimental					
11	F	26	15	32	31
12	M	30	18	34	36
13	M	26	22	41	31
14	M	30	12	37	31
15	F	41	18	39	44
16	M	27	16	32	33
17	M	22	17	28	35
18	M	31	19	26	32
19	F	18	10	27	15
20	M	30	18	27	29

Student number	Sex	DS score	RPM pretest	RPM posttest	RPM retention test
White - high DS - Hawthorne control					
21	M	50	15	20	22
22	M	74	17	18	16
23	M	50	14	17	20
24	M	81	17	16	14
25	F	68	19	26	26
White - low DS - Hawthorne control					
26	M	40	19	20	13
27	M	36	22	21	20
28	M	31	14	14	7
29	F	40	16	26	35
30	F	36	19	20	30
White - high DS - pure control					
31	M	62	23	31	26
32	M	51	18	17	12
33	F	50	18	25	20
34	F	81	15	19	18
35	M	46	24	33	39
White - low DS - pure control					
36	M	14	25	30	28
37	F	40	16	17	17
38	F	40	18	15	9
39	M	26	21	20	17
40	M	19	18	22	12

Student number	Sex	DS score	RPM pretest	RPM posttest	RPM retention test
Black - high DS - experimental					
1	F	74	12	31	31
2	M	85	20	31	35
3	M	81+	11	35	36
4	F	50	13	33	33
5	F	68	17	33	33
6	F	68	13	28	34
7	F	50	12	25	25
8	M	51	14	42	43
9	M	51	17	30	30
10	M	62	18	32	34
Black - low DS - experimental					
11	M	35	13	14	26
12	M	26	13	16	15
13	F	27	13	16	16
14	M	26	6	20	16
15	M	40	14	36	41
16	M	26	18	28	36
17	M	26	14	17	13
18	F	35	21	24	24
19	F	31	17	20	18
20	M	30	14	17	22

Student number	Sex	DS score	RPM pretest	RPM posttest	RPM retention test
Black - high DS - Hawthorne control					
21	M	50	13	11	9
22	F	52	10	7	10
23	F	68	13	12	14
24	F	57	14	11	12
25	F	50	11	9	13
Black - low DS - Hawthorne control					
26	M	40	15	15	15
27	M	36	18	14	26
28	F	35	15	9	9
29	F	22	21	26	15
30	F	26	12	13	21
Black - high DS - pure control					
31	F	89	15	18	22
32	M	76+	20	23	37
33	M	81	12	7	9
34	M	63	16	15	18
35	F	51	18	16	19
Black - low DS - pure control					
36	M	26	13	10	18
37	F	26	19	23	23
38	F	35	19	27	43
39	M	22	15	15	11
40	M	40	15	12	17