R E P O R T   R E S U M E S

ED 020 787 PS 000 987

THE JOHNS HOPKINS PERCEPTUAL TEST, THE DEVELOPMENT OF A RAPID INTELLIGENCE TEST FOR THE PRE-SCHOOL CHILD.

BY- ROSENBERG, LEON A. AND OTHERS
JOHNS HOPKINS UNIV., BALTIMORE, MD.

REPORT NUMBER OEO-510 PUB DATE APR 66
EDRS PRICE MF-$0.25 HC-$0.76 17P.

DESCRIPTORS- *TEST CONSTRUCTION, TEST VALIDITY, *PRESCHOOL TESTS, *INTELLIGENCE TESTS, *CULTURE FREE TESTS, *PERCEPTION TESTS, INFORMATION THEORY, INTELLIGENCE QUOTIENT, RETARDED CHILDREN, PILOT PROJECTS, STANFORD BINET, PPVT, DAP, CMMS,

IN ORDER TO DEVELOP AN INTELLIGENCE TEST FOR PRESCHOOL-AGE CHILDREN THAT WOULD OVERCOME SOME OF THE LIMITATIONS OF AVAILABLE TESTS, A PERCEPTUAL DISCRIMINATION TEST USING POLYGONAL FORMS HAS BEEN DESIGNED AND TESTED. THE CHILD POINTS TO ONE OF TWO, THREE, OR FIVE FORMS MATCHING A STIMULUS FORM. INITIAL TESTING WITH 44 CHILDREN RANGING IN AGE FROM 3 YEARS TO 5 YEARS 10 MONTHS SHOWED GOOD CORRELATION OF THE TEST SCORES WITH AGE AND THE RAW SCORES OF THE PEABODY PICTURE VOCABULARY TEST (PPVT). A SHORTENED TEST HAVING 30 CHOICES WAS DEVELOPED AND GIVEN TO TWENTY-TWO 4- TO 5-YEAR-OLD NURSERY SCHOOL CHILDREN. AGAIN, GOOD CORRELATIONS WITH AGE, THE PPVT, AND THE DRAW-A-PICTURE (DAP) TEST WERE OBTAINED. VALIDITY STUDIES OF THIS TEST, DESIGNATED THE JOHNS HOPKINS PERCEPTION TEST (JHPT), HAVE SO FAR INVOLVED SOME 340 CHILDREN. STATISTICALLY SIGNIFICANT CORRELATIONS OF THE SCORES ON THE JHPT WITH SCORES ON THE PPVT, THE DAP, AND THE COLUMBIA MENTAL MATURITY SCALE (CMMS) HAVE BEEN OBTAINED. A COMPARISON WAS MADE OF THE PERFORMANCE OF 37 UPPER AND 52 LOWER SOCIOECONOMIC STATUS CHILDREN ON THE JHPT, THE PPVT, AND THE DAP. THE DIFFERENCES IN MEAN SCORES BETWEEN GROUPS WAS HIGHEST FOR THE PPVT, LESS MARKED FOR THE DAP, AND STATISTICALLY INSIGNIFICANT FOR THE JHPT, SUGGESTING THAT IT MAY BE A CULTURE-FREE INSTRUMENT. THIS PAPER WAS PRESENTED AT THE EASTERN PSYCHOLOGICAL ASSOCIATION ANNUAL MEETING (NEW YORK, APRIL, 1966). (DR)

FILMED FROM BEST AVAILABLE COPY

ED020787
PS 000987
E

U. S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

The Johns Hopkins Perceptual Test

The Development of a Rapid Intelligence Test for the Pre-School Child[1,2,3]

Leon A. Rosenberg, Ph.D., Anna M. Rosenberg, Ph.D.,
and
Michael Stroud, B.S.

The Johns Hopkins University School of Medicine

1 Supported in part by the Office of Economic Opportunity (Contract #510; principal investigator: Leon Eisenberg, M.D.).

2 Supported in part by the Children's Bureau, Department of Health, Education & Welfare (Grant # H-135; principal investigator: Leon A. Rosenberg, Ph.D.).

3 Presented at the Eastern Psychological Association Annual Meetings, April, 1966, New York, N.Y.

ERIC

15839

Available tests of intellectual function have serious limitations when applied to: (1) children who do not speak because of functional or organic handicaps; (2) culturally deprived children with limited verbal and experiential repertoires; (3) children with motor handicaps; (4) very young or retarded children. These limitations lead to considerable difficulty in the important diagnostic task of identifying the handicapped pre-school child and discriminating between the defective and the "pseudo-defective" youngster. Thus, the designation of an appropriate therapeutic regime may be delayed (and at times inappropriate management provided).

The purpose of this research was to develop and evaluate the usefulness of a new technique for the intellectual assessment of pre-school age children. Our purpose was to develop an instrument for the evaluation of young children in whom a variety of difficulties impair performance on available tests. Among the general pediatric population, for example, are youngsters who do not speak for either functional or organic reasons and one is, therefore, limited in regard to what instruments can be used. Similar difficulties occur when one is dealing with the culturally deprived youngster with limited verbal abilities. Shifting to "performance" tasks when faced with this difficulty is a typical solution which, however, is often not successful when one is dealing with the quite young child, or the immature youngster with slow physical development, or the organically damaged youngster with impaired visual-motor coordination or the retarded youngster. The child may be unable to perform on any of the available tasks because they are too difficult for his level of maturity or for his degree of retardation, or because of impaired motor skills. In these cases, for example, the Bender Visual-Motor Gestalt Test is reduced to scribbling and the items on the Stanford-Binet Intelligence Test are often too difficult at even the youngest age levels (e.g. the spastic child cannot manipulate the Binet formboard or balance the blocks). The available formboard tests are often too simple and one becomes dependent on time measures which are quite variable and for which there are inadequate norms. Furthermore, formboard tests cannot be used with many handicapped children. A test like the Draw-A-Person (DAP) also gives highly variable results and depends too much on the child's willingness to manipulate a pencil. The drawback, with children below age 6, in regard to the Bender-Gestalt and the DAP, is not only in terms of drawing ability but, of even more importance, in terms of the child's willingness to draw. Many youngsters, for one reason or another, approach the Bender task or the DAP with statements such as "I can't draw - I'm no good at that." A task depending simply on a drawing response appears to be quite limited when we are dealing with youngsters who are either too young to draw, or who are negativistic; or who are too immature or retarded to be able to do more than scribble.

What we felt was needed was a task which could be ordered along a well-defined continuum of complexity; a series which would reach the low level needed for young and/or retarded youngsters and would also be adequately sensitive to the range of individual differences at even the youngest age level. Many of the problems of examining this kind of child would be resolved if, at each level of complexity, there were more than one way of evaluating the youngster's performance. For example, tasks could be developed which utilized different responses such as pointing to the object, fitting the

15841

object into a formboard, and copying the object by drawing. The advantage would lie in the ability to shift to an equivalent task requiring a different response when a specific handicap interferes with a particular response. An important advantage from the diagnostic point of view would be the resulting comparison of performance on these three different types of tasks. One might thus demonstrate difficulties in visual-motor coordination by a comparison between a task requiring a pointing response with a task requiring formboard fitting or copying. If such tasks requiring different responses could be made equivalent in quantitative terms, then such comparisons could be made successfully.

What we felt these problems required was a series of tasks using different responses, which could be successfully administered to young children in spite of the presence of a variety of functional or organic handicaps. The problem was one of finding a means of defining complexity by which a series of tasks along a well-defined continuum of complexity could be designed. An information theory approach held the best promise of success in the construction of such tasks. Our decision to utilize simple perceptual designs in what would basically be a discrimination task posed several problems. First, how to choose the pattern or shape of the design. To reduce the effect of prior learning, these designs should be "random" and as "culture-free" as possible. Secondly, the designs had to be ordered along a continuum of complexity by some rational means. The most promising procedure for solving these two problems was the application of information theory to shape and pattern perception presented by Attneave and Arnouit (1956). The procedure has two facets which are directly related

ERIC

15842

to the questions raised: it allows for the generating of perceptual patterns or shapes on a completely random basis (i.e. the experimenter has no control over what shape the figure will assume); it permits the generation of an unlimited series of figures from the same "stimulus-domain" (i.e. the figures are a sample of stimuli from a parent population characterized by certain determinant statistical parameters). Changing the stimulus population by known variation of one or more parameters produces a different series of figures. The procedure for producing different stimulus populations can be seen as increasing the informational content of the individual figures. Hence, different samples of figures or shapes can be ordered along a continuum of complexity if one defines complexity as the amount of information (or uncertainty) "built-in" to such a perceptual figure.

The use of information measures produces a second approach to the problem of defining the difficulty level of the task. The amount of information (or the degree of uncertainty) in a choice-making task is measured by the formula: $U = \log_2 k$; where U = uncertainty of information expressed in bits, and k = the number of equally probable alternatives (Garner, 1962). Hence, the difficulty level of the total task could be defined in terms of the total number of alternatives which make up the task.

The Attneave and Arnoult process produces a series of random designs, the shapes of which will vary greatly. Figure 1 demonstrates the procedure involved.

It is also possible to generate a series of designs that systematically vary along a known continuum. The procedure, developed by LaBerge and Lawrence (1955), involves first producing a random shape by the Attneave method and then assigning

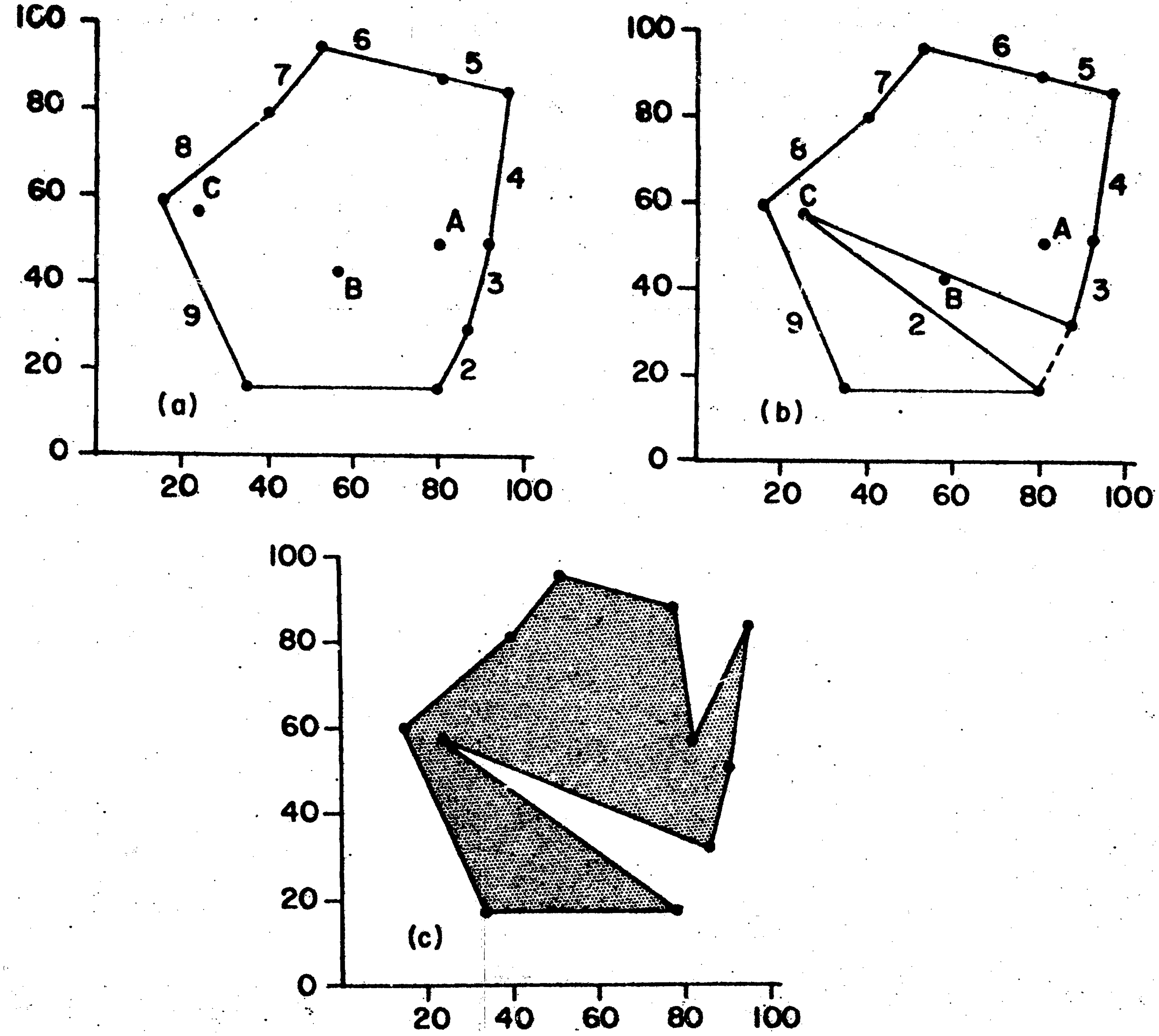

FIG. I. SUCCESSIVE STAGES IN THE CONSTRUCTION OF A "RANDOM" FIGURE

15843

15844

each point on the contour randomly chosen "X" and "Y" increments to its coordinates. These new coordinates are plotted and connected as a new matrix. The same increments are then added to the new coordinates and yet another figure is constructed. This process will produce a row of figures each differing from the adjacent figure by a constant amount of distortion as measured by the distance through which the points move. Any two adjacent shapes in such a row are equally spaced in terms of the average distance the points have moved.

It was decided to utilize both of these procedures to produce the needed perceptual figures. Hence, there were two types of tasks; one involving purely random figures and one using figures related to each other along a continuum of known variation.

The complexity of the task was defined in two ways: (1) The amount of information in the design itself. This was defined as the number of turns in the contour of the shape. In this case, it was the number of angles in the shape. This procedure follows the work of Attneave (1957) who demonstrated that "number of turns" was the most important variable in determining the judged complexity of random shapes. (2) The number of alternatives in the discrimination task. If the subject had to choose between two figures in trying to find a match for a stimulus figure, the degree of uncertainty in that task can be expressed as one bit (or one can say that the task entails one bit of information). A task with five choices, would be a 2.3219 bit problem. The more bits of information presented to the subject, the more complex the task.

We were interested in three types of tasks: (1) a task where the child only had to point; (2) a task where he had to fit a form into a formboard; (3) a drawing task.

15845

Our work up to the present has been concerned with the first task only and the present paper presents the instrument that was developed.

Two types of designs were utilized: Type A - a series of random designs; Type B - a series of designs varying by known increments from an original design. Types A and B designs were developed at three different information levels; designs with three points (a point being a "turn" or change in contour), four points, and six points. The task for the child was to point to the design which matched the stimulus design. The youngsters were faced with either two, three, or five alternatives. Hence, the study examined the variables of a number of alternatives, complexity of the designs, Type A versus Type B, on accuracy in a perceptual matching task.

The subjects were 44 children, who ranged in age from 3 years-0 months, to 5 years-10 months. The occupations of their parents ranged from semi-skilled hospital employees to professional staff. The Peabody Picture Vocabulary Test (PPVT) was also administered to the children. The results indicated that degree of accuracy on the perceptual task was highly correlated with both age ($r = .618$) and the PFVT raw score (.744). Partial correlation indicated that the major relationship was between the PPVT and perceptual accuracy.

From this data, a shorter task was chosen where a youngster would have to make a series of 30 discriminations calling for a total time expenditure of 15 minutes. The new task was administered to 22 four-to-five year-old youngsters enrolled in a middle-class private nursery school. The PPVT and DAP were also administered. The data obtained indicated a step-wise increase in difficulty level from one task to

ERIC

15840

the next, and correlations with the PPVT of .566, with the Harris (1963) scoring of the DAP of .548, and with age of .584.

It appeared, then, that we had developed a simple, pleasant (e.g. no child at even the youngest age refused the task, although many were reluctant to do the DAP and all children considered the task as "fun", "a good game", etc.), and rapid test of intellectual functioning. The uniqueness and "randomness" of the designs suggested that the task might be relatively culture-free. The lower level of the task appeared simple enough for the very young or very retarded. The response required of the subject involved a very simple motor act that could be performed by physically handicapped children. The instructions might well be successfully communicated in pantomime for the testing of deaf children.

Our validity studies, up to this point, have utilized some 800 children. These youngsters ranged in age from 3 to 6 years and were obtained from private nursery schools and a county well-baby clinic. Table 1 presents Pearson correlations of the Johns Hopkins Perceptual Test (JHPT) with the PPVT, DAP, and Columbia Mental Maturity Scale (CMMS). All correlations were statistically significant. For each test the correlations were lower in the lower-socioeconomic group. This was caused by a depression of their scores on the criterion measures. It was this data that indicated that the Johns Hopkins Perceptual Test was indeed measuring some aspect of intelligence.

We were also able to compare upper and lower-socioeconomic status children in terms of their performance on the JHPT, the PPVT, and the DAP. Figure 2 presents

TABLE 1

Correlational Data

Middle-Socio-Economic Group		
JHPT x PPVT	:	.515, $p < .01$, $N = 50$
x DAP	:	.702, $p < .01$, $N = 37$
x CMMS	:	.798, $p < .01$, $N = 25$
Lower Socio-Economic Group		
JHPT x PPVT	:	.449, $p < .01$, $N = 79$
x DAP	:	.356, $p < .05$, $N = 36$
x CMMS	:	.657, $p < .01$, $N = 78$

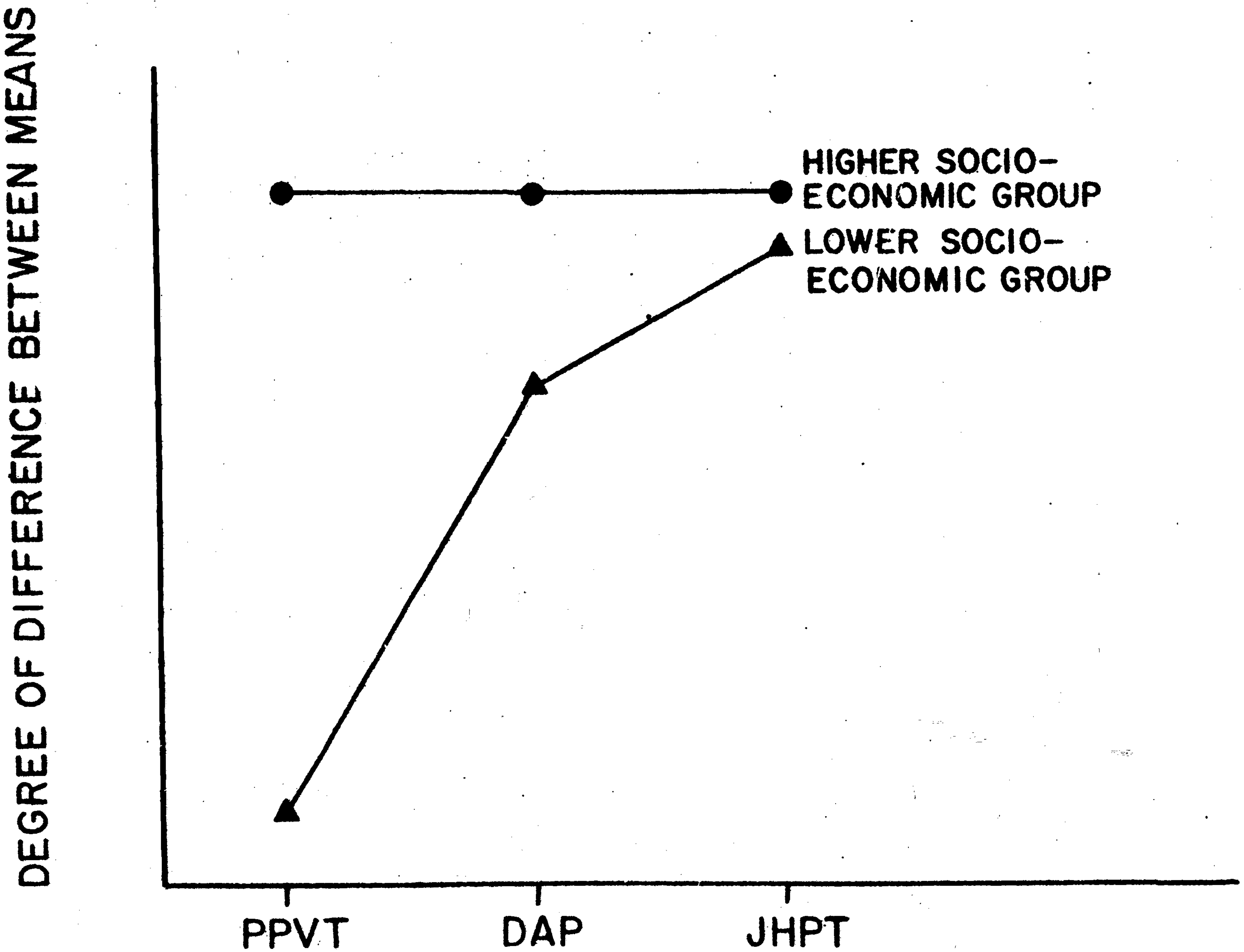

FIG. 2. SIZE OF DIFFERENCE BETWEEN MEAN SCORES ON THREE TESTS OF INTELLIGENCE WHEN COMPARING HIGHER SEC CHILDREN WITH LOWER SEC CHILDREN

15849

the figures graphically. There were 37 children in the group titled "higher socio-economic group" and 52 subjects in the group titled "lower socioeconomic group". Figure 2 indicates the size of the difference between mean scores on the three tests of intelligence. It was quite apparent that there was an extreme difference on the Peabody, less difference on the DAP, and much less difference on the JHPT. Table 2 presents the statistical tests of this data. The analysis demonstrated that the difference on the PPVT and the difference on the DAP were statistically significant while there was no significance to the obtained difference on the JHPT. This data raised the possibility that the JHPT might well be a relatively culture-fair instrument.

One additional study has been completed which is of interest in terms of the practical application of the JHPT. This particular study was initiated as the result of our finding a significant difference in the prevalence of retardation when one compared the PPVT and CMMS to results obtained with the JHPT. All three tests had been administered to a group of kindergarten-age children. The PPVT estimated the incidence of retardation (I.Q. equal to or below 69) to be 39.6%; the CMMS, 17.9%; the JHPT, 3.5%. Since the JHPT was and is still in an experimental stage it obviously could not be used as a criterion measure. Hence, the Stanford-Binet Intelligence Scale was administered to each child.

The subjects were 28 Negro children enrolled in a slum-area kindergarten. There were 16 males and 12 females with an average age of 62.94 months and an age range of 50 to 70 months. The Stanford-Binet agreed with the JHPT in indicating that there were 3.5% retarded youngsters. The study gave us some additional data in regard

15850

TABLE 2

Middle Class versus Lower Class

Test	Difference Between Raw Score Means	"t"		
PPVT	9.72	4.24	p	.0025
DAP	3.16	2.49	p	.025
JHPT	.91	n.s.		

ERIC

to the PPVT and CMMS. The mean I.Q. score obtained on the PPVT was 86.52; on the CMMS, it was 83.96; while the Stanford-Binet yielded a mean I.Q. of 91.32. The differences between the PPVT and Stanford-Binet and between the CMMS and Stanford-Binet were statistically significant (p .01).

Our data indicated that the youngsters labelled "mental defective" by the PPVT and CMMS were not simply children of borderline I.Q. who moved a bit lower on these two tests. They also included children whose intelligence, according to the Stanford-Binet, fell in the 90 to 109 range.

In evaluating this data it should be kept in mind that the JHPT takes 10 to 15 minutes to administer and does not require any extensive training or experience. The Stanford-Binet, on the other hand, certainly requires a good deal more time and a great deal more sophistication on the part of the examiner. Although both tests agreed in regard to the prevalence of retardation, they did not agree as to exactly which child was retarded. The retarded child identified by the JHPT had an I.Q. in the 70 range on the Stanford-Binet. The retarded youngster identified by the Stanford-Binet did not obtain the lowest score on the JHPT but did score in the lowest quartile. It should also be kept in mind that this data does not imply that the Stanford-Binet is culture-fair. Although very careful administration of the Stanford-Binet may not overestimate the prevalence of retardation with disadvantaged children, the test still underestimates the percentage of youngsters in the upper ranges of intelligence. Hence, the distribution of Stanford-Binet scores is extremely constricted while the distribution of JHPT scores more closely followed the expected normal curve for intelligence measures.

Our current research is moving in several area. It appears that the ceiling of the test is too low. Too many youngsters of normal intelligence at age 6 and 7 receive the same score. We are, therefore, increasing the length of the test by adding more complex design series. The time involved in the administration of the extended version of the test would be no more than 15 to 20 minutes.

We are currently obtaining data on children at different levels of retardation. A sample of 30 institutionalized retarded youngsters ranging in age from 5 to 9 years-old were administered both the JHPT and the Stanford-Binet. The correlation between the JHPT raw score and the Stanford-Binet mental age was .70. A large sample of minimally retarded 6 and 7 year-olds is currently being examined through the help of the Baltimore County Board of Education, Department of Special Education. Additional data in the area of the prediction of retardation will become available when the entire population of 4 year-olds within one county in the state has been tested and followed for some time.

We are also examining the possibility of improving our scoring system. At present, the JHPT raw score is simply the sum of all correct responses. We are moving toward a weighted scoring system which will take into account differences in difficulty level of the individual tasks.

An intriguing study which will yield data both in the area of retardation and cultural-influences will soon be under way involving two groups of Nigerian 4 and 5 year-olds. One group will have a much higher incidence of retardation than the

ERIC

other, and both will be tested with the JHPT and Porteus Maze Tests. And finally, in the area of physical handicaps, samples of brain-damaged youngsters and youngsters with different degrees of deafness will soon be added to our series of studies.

ERIC

References

Attneave, F. Physical determinants of the judged complexity of shapes. J. exp. Psychol., 1957, 53, 221-227.

Attneave F., and Arnoult, M.D. The Quantitative Study of Shape and Pattern Perception. Psychol. Bull., 1956, 53, 452-471.

Garner, W.R. Uncertainty and Structure as Psychological Concepts. 1962, John Wiley, Inc., New York.

Harris, D.B. Children's Drawings as Measures of Intellectual Maturity. 1963, Harcourt, Brace and World, Inc., New York.

LaBerge, D.L. and Lawrence, D.H. A Method of Generating Visual Forms of Graded Similarity. Amer. Psychologist, 1955, 10, 401.