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

ED 014 754 CG 000 859

THE JOHNS HOPKINS PERCEPTUAL TEST--ITS DEVELOPMENT AND CURRENT STATUS AS A MEASURE OF INTELLECTUAL FUNCTIONING.

BY- ROSENBERG, LEON A.

PUB DATE 66

EDRS PRICE MF-$0.25 HC-$1.00 23P.

DESCRIPTORS- *TEST CONSTRUCTION, *INTELLIGENCE TESTS, *MEASUREMENT, RESEARCH, *CULTURE FREE TESTS, TEST VALIDITY, TEST RELIABILITY, PERCEPTION TESTS, *PERFORMANCE TESTS, CORRELATION, JOHNS HOPKINS PERCEPTUAL TEST, PEABODY PICTURE VOCAB. TEST, DRAW A PERSON, COLUMBIA MENTAL MATURITY SCALE

THE JOHNS HOPKINS PERCEPTUAL TEST (JHPT), REVIEWED IN THIS PAPER, WAS DESIGNED TO MEET THE NEED FOR A BRIEF, EASILY ADMINISTERED, NON-VERBAL INSTRUMENT TO ASSESS THE INTELLECTUAL FUNCTIONING OF CHILDREN. THE DATA ON THIS INSTRUMENT INDICATES THAT IT IS RELATIVELY CULTURE FAIR AND USEFUL FOR TESTING CHILDREN WHOSE PHYSICAL HANDICAPS PRECLUDE THE USE OF STANDARD INSTRUMENTS. FOLLOWING A REVIEW OF THE RATIONALE FOR THE DEVELOPMENT OF A CULTURE FAIR INSTRUMENT, THE DEVELOPMENT OF THE JHPT IS DISCUSSED. IT WAS FOUND TO BE A REASONABLY RELIABLE AND VALID TEST OF INTELLIGENCE WHICH CAN BE ADMINISTERED BY A NON-PROFESSIONAL TO YOUNG CHILDREN. THE ISSUE OF CULTURE FAIRNESS IS NOT SO EASILY DEALT WITH AND FURTHER RESEARCH IS NEEDED. PRESENT STUDIES UTILIZING THE JHPT ARE BEING CONDUCTED BOTH IN THE UNITED STATES AND ABROAD. THE NEED FOR CULTURE FAIR INSTRUMENTS IS EXPLORED. THIS PAPER WILL BE PUBLISHED AS A CHAPTER IN THE DISADVANTAGED CHILD, VOLUME 2, SPECIAL CHILD PUBLICATIONS, SEATTLE, WASHINGTON. (SK)

Pre-publication draft: to be published as a chapter in:
The Disadvantaged Child, Vol. 2, Special Child
Publications, Seattle, Wash.

The Johns Hopkins Perceptual Test: Its Development and Current Status as a Measure of Intellectual Functioning[1, 2]

on A. Rosenberg, Ph.D.

Assistant Professor of Pediatrics and Medical Psychology
The Johns Hopkins University School of Medicine
and
Pediatric Psychologist
The Johns Hopkins Hospital

[1] Part of the data discussed in this chapter has been presented in the following: (1) Rosenberg, L.A., Rosenberg, Anna M., and Stroud, M. The Johns Hopkins Perceptual Test: the development of a rapid intelligence test for the preschool child. Presented at annual meetings of Eastern Psychol. Assoc., New York, N.Y., April, 1966; (2) Rosenberg, L.A. Identifying the "gifted child" in the culturally deprived population - the need for culture-fair instruments. Amer. Jour. Orthopsychiat., 1967, 37, 342-343.

[2] Supported in part by OEO contract number 510 and grants H-135 and H-135 (C1) from the Children's Bureau of the Department of Health, Education, and Welfare.

The Johns Hopkins Perceptual Test was developed in response to a need for a brief, simple to administer, and non-verbal test of intellectual functioning for young children. The data to be presented suggests that the instrument might be relatively "culture-fair" and that the instrument might be of use with a variety of children where physical handicaps preclude the use of standard instruments. The following review will make it quite clear that we are in the midst of examining the usefulness and limitations of this test and that the instrument is still in an early stage of development.

This chapter will follow the following outline: A brief statement of our rationale for the development of a culture-fair instrument; the development of the Johns Hopkins Perceptual Test (JHPT) and our current data; a description of ongoing and future research; a statement of our expectations regarding the practical use of the JHPT; and a statement of the author's personal prejudice in regard to the intellectual evaluation of disadvantaged children.

The rationale we have been following for the development of culture-fair instruments involves two steps. First, does the instrument in question measure "intelligence." The usual criterion measures involve correlations with (1) school performance and (2) "established" intelligence tests. To validate a culture-fair instrument requires a restriction n the sample of subjects used. Since disadvantaged children tend to demonstrate less-than-adequate school performance and to perform poorly on current intelligence tests, correlations between a culture-fair test and these two criterion measures would be expected to be quite low due to the restriction in range of the lower socio-economic status (SES) subjects' scores on the criterion measures. Hence, our validity studies done with the JHPT involved both low and middle SES groups, with the correlations obtained with the more culturally advantaged group being the actu . basis for statements of validity. Step two, in our methodology, moves on to the question of "culture-fairness." Our definition of "fairness" is a narrow one: a test is "culture-fair" if the obtained scores are demonstrated to be free of the influence of SES differences between subjects; the distributions of test scores obtained from groups of subjects representing the extreme points of the SES continuum should be identical.

Our original objective was to produce a measure of intellectual functioning which would be of value when one was evaluating large numbers of young children. We were concerned primarily with pre-school and early school years (ages 3-0 to 7-11 constitute the maximum age range dealt with presently). The instrument had to meet several criteria:

(1) The test had to call for a simple response which would allow for the successful testing of mute, physically handicapped, or somewhat uncooperative children who might balk if asked to perform a complex act.

(2) The instrument would have to be simple to administer and score and should involve approximately 20 minutes of the child's time. The need for speed being related to our interest in producing a tool useful in large-scale testing programs but being an individually administered test. This approach was based on a belief that young children and low SES children do poorly on group administered tests. The instrument had to be simple to allow for non-professional technicians to successfully use it. This attribute would make the tool useful in situations where many non-psychologists would be used as testers.

(3) The items of the test should have some promise in regard to being "culturally fair."

We decided to utilize simple perceptual designs in what would basically be a discrimination task. The development of such a discrimination task poses 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-fair" as possible. Secondly, we felt that the designs had to be ordered along a continuum of complexity by some rational means. We felt that the most promising procedure for solving these two problems was the application of information theory to shape and pattern perception presented by Attneave and Arnoult (1956). The procedures developed by these authors had two facets which were directly related to the questions raised. It allowed for the generation of perceptual patterns or shapes on a completely random basis (i.e. the experimenter had no control over what shape the figure would assume); it permitted the generation of an unlimited series of figures from the same "stimulus-domain" (i.e. the figures would be 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 would produce a different series of figures. The procedure for producing different stimulus populations could be seen as increasing the informational content of the individual figures. Hence, different samples of figures or shapes could be ordered along a continuum of complexity if one defined complexity as the amount of information (or uncertainty) "built-in" to such a perceptual figure.

The use of information measures would also produce 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.

Briefly, the Attneave and Arnoult process allows one to decide on the number of points or angles that a figure will have. From that point on, the exact shape that the figure will assume is beyond the experimenter's control. Square grid coordinate paper is utilized with the X and Y coordinates of each point determined by a table of random numbers. A well-defined procedure is then followed for the connecting of those points. For a better understanding of the process the reader is referred to the Attneave and Arnoult article.

The process produces a series of random designs, the shapes of which will vary greatly. 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), requires one to first produce a random shape by the Attneave and Arnoult method and then assign 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 total 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 the stimulus figure, the degree of uncertainty in that task could 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.

Our earliest research involved these two types of designs which were labelled as follows: 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 2, 3, or 5 alternatives. Hence, the study examined the variables of number of alternatives, complexity of the designs, type A versus type B, on accuracy in a perceptual matching task.

In this first study, 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 the degree of accuracy on the perceptual task was highly correlated with both age (r = .618) and the PPVT 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 Draw-A-Person (DAP) were also administered. The data obtained indicated a step-wise increase in difficulty level within the JHPT, and correlations with the PPVT of .566, with the Harris (1963) scoring of the DAP of .648, 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-fair. The lower level of the task appeared simple enough for the very young or even the 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.

Before going on to discuss reliability and validity data, we will take a moment to describe our test procedure. Figure 1 presents a typical test situation. The child has been presented with five designs lying flat on a board. Designs which match one of the five are presented one at a time and are placed upright on the stand shown in the top center of Figure 1. As shown in the figure, the child then points to the one card in the display of five which he feels matches the upright figure. Although there are some variations, this is the basic procedure followed in the administration of the JHPT. Before presenting these materials to the child, we establish an appropriate "set" by utilizing

FIGURE 1

The Administration of the Johns Hopkins Perceptual Test plastic triangles, circles, and squares. The child is given one of each of these figures and the examiner then holds up an identical circle, square, or triangle. He asks the child to find which one of his forms matches the one held in the examiner's hand. The use of 3-dimensional forms allows the examiner to aid the child by placing the child's forms next to his own. These 3-dimensional forms are used until the examiner is certain that the child understands that a matching task is what he is dealing with. The 3-dimensional forms are then placed out of sight and formal testing begins. Each of the test items are photographs consisting of a black figure on a white background. Each photograph is mounted on heavy cardboard and is 5 inches by 7 inches in size.

A series of studies have been completed which answer the questions of reliability, validity, and "culture-fairness." The validity data is presented in Table 1. As expected, the correlations are lower for the "low" SES groups. Some additional validity data is also available. A sample of 30 institutionalized retarded children (age range of 5 to 9, mean I.Q. of 51.60, S.D. = 14.47) were administered both the JHPT and Stanford-Binet. The correlation between the JHPT raw scores and the mental ages obtained with the Stanford-Binet was .697 (there were two subjects who were not included in the study. Both could not respond to the JHPT. One had a Binet I.Q. of 42 and the other subject's Binet was "unscoreable"). Combining this group of 30 with 57 kindergarten and first grade students (a group attending a "slum-area" school whose Stanford-Binet I.Q. scores "piled up" in the "Borderline" and "Dull Normal" I.Q. ranges) we obtained an r of .572 ($p < .01$) between Binet M.A. and JHPT. Partialling-out the age variable, we obtained an r of .501 ($p < .01$) between Binet I.Q. and JHPT.

A reliability study has also been completed and has yielded data both in regard to test-retest reliability and to the degree to which one can use non-professional examiners

TABLE 1

Validity Coefficients[1,2]

		r	N
(1965) "Middle" SES (Preschool)			
JHPT	x Peabody Picture Vocabulary Test:	.615**	50
	x Draw-A-Person	.702**	37
	x Columbia Mental Maturity Scale:	.798**	25
"Low" SES (Preschool)			
JHPT	x Peabody Picture Vocabulary Test:	.449**	79
	x Draw-A-Person	.356*	36
	x Columbia Mental Maturity Scale:	.657**	78
(1966) "Middle" SES (Preschool)			
JHPT	x Peabody Picture Vocabulary Test:	.756**	71
	x Columbia Mental Maturity Scale:	.761**	71

* = $p < .05$ ** = $p < .01$

[1] All correlations are based on raw scores.

[2] The SES designations of "middle" and "low" are only rough indicators. The "middle" group were children attending private nursery schools while the "low" group were obtained from county Well-Baby clinics with many families being on welfare.

with the JHPT. Two examiners were utilized. One was a 24 year-old female college graduate with no graduate level training. The second was a 15 year-old female high school student. The results are presented in Table 2 and indicate satisfactory reliability. The data suggests that more satisfactory results are obtained when the examiner is an older college level person.

Thus far, the data appears to indicate that JHPT is a reasonably reliable and valid brief-test of intelligence which can be successfully administered to young children by a non-professional examiner. The validity data is limited to correlations with other tests of intelligence and no attempt has been made to predict school performance.

The assertion that the JHPT is a "culture-fair" instrument is based on one small study. Two groups of children of different SES levels were compared in terms of their performance on the JHPT, the PPVT, and the DAP. The two groups were given the labels of "lower" and "middle" as descriptions of their social status. The labels imply difference only and do not mean that the social status of each child's family was actually determined. The "lower" group consisted of 52 pre-school children who were attending a county well-baby clinic. The clinic served an impoverished area and many of the families were receiving welfare assistance. The "middle" group consisted of 52 pre-school children attending a private nursery school which served a "middle-class" section of Baltimore City. The fact that the two groups differed in social status was quite clear, although no attempt was made to accurately determine the actual SES level of each group in terms of local class structure.

The two groups had equal mean ages and ranged in age from 4 to 5. Sex and race were uncontrolled with the race difference being most blatant. The "lower" group were all Negro while the "middle" group were all white. The data obtained from this hastily

TABLE 2

Test-Retest Reliability Coefficients for Three

Tests with a Time Lapse of 17 Days Between Test Sessions

Examiner	N	Test Peabody Picture Vocabulary	Columbia Mental Maturity	JHPT
15 year-old	22	.889	.894	.864
24 year-old	39	.886	.935	.933
Total	61	.888	.917	.902

Mean Test-Retest Difference in Raw Scores

Examiner	PPVT	CMMS	JHPT
15 year-old	3.81	5.55	2.71
24 year-old	1.24	7.30	- 0.36
Total	2.15	6.64	0.76

done and inadequately controlled study was quite promising. Figure 2 shows the distributions (ogives) of the raw scores obtained on the three tests. It is quite apparent that there was an extreme difference on the PPVT, less difference on the DAP, and much less difference on the JHPT. Table 3 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 is the data that raised the possibility that the JHPT might well be a relatively culture-fair instrument.

Although promising, the JHPT itself and the research data discussed have several limitations. Just what aspect of intellectual functioning does the test measure? Although it only requires matching of designs, the JHPT correlates significantly with measures involving widely different types of intellectual activity; from the concept-formation task of the Columbia Mental Maturity Scale to hearing-vocabulary in the PPVT to the varied skills sampled by the Stanford-Binet. We could look to the "g" factor in theories of intelligence for an answer but this question would be best resolved by a large scale factor-analytic study. At present, the exact answer to this question remains unavailable.

Is the present data sufficient to allow us to claim that the instrument is definitely "culture-fair" within the limits of our definition of that term? An analysis of the available data must lead to an answer of "no."

The study discussed above suffers from the following: (1) Age range sampled was much too limited; (2) Numbers of subjects in each group were too small; (3) The means used to establish that the two groups were different in SES level were too poorly defined and crudely administered leaving open the possibility that some unknown number of subjects might have been misidentified as being of "low" or "middle" SES; (4) The

FIGURE 2

Distributions (Ogives) of raw scores obtained with three tests of intelligence administered to two different socio-economic class groups

TABLE 3

Tests of the Statistical Significance of the Differences in Raw Score Means Obtained on Three Tests of Intelligence Administered to Two SES Groups

Test	Size of Difference	"t"	
PPVT	9.72	4.24	$p < .0025$
DAP	3.16	2.49	$p < .025$
JHPT	.91	less than 1	

obtained data itself suggests the need for further research. The ogives of Figure 2 demonstrate not only the potential "fairness" of the JHPT as compared to the DAP and PPVT, but also a possible degree of difference, or "unfairness," between the two groups at the upper range of JHPT raw scores. At the 25th percentile, the difference between the raw scores of the two groups was .50; at the 50th percentile, it was .75; at the 75th percentile, it was 1.35; and at the 90th percentile, it was 2.50. Although no statistically significant difference was found in regard to differences between means, there is the suggestion that some SES influence might have effected the scores at the upper levels of the JHPT.

Hence, we can conclude that although the JHPT appears to be a promising tool, there is a great need for further research. Several interesting studies are in progress at this time. Samples of minimally brain-damaged children are being tested with an elaborate psychological battery. The JHPT has been included and this data will give us excellent insights into the effect of CNS dysfunction on performance on the JHPT. A large number of children between the ages of 3 and 9 who suffer from various degrees of hearing impairment are being tested with both the JHPT and the Leiter International Performance Scale. On a smaller group of 30, we obtained an r of .845 between the JHPT and the Leiter. This larger series of subjects will enable us to determine the value of the JHPT in the examination of deaf children.

The JHPT, along with other psychological tests, is being used in research studies in other countries by several people. Data is currently being collected in Nigeria and Guatemala and plans are being made to use the instrument in Saudi Arabia. In all three studies the subjects will be pre-school age children and large group comparisons will be made. Two studies in this country are under way which will give us an opportunity to do long term follow-up in regard to achievement differences between groups scoring high and low on the JHPT. In both studies the JHPT is included in a battery of tests which will allow us to examine the relationship between the JHPT and these other tests. In one situation, the entire first grade of a racially-integrated poverty-area elementary school has been tested and these youngsters will be followed for some time. In the second situation, a large number of children attending a Head-Start program will be tested and followed through later school years. In regard to the question of culture-fairness, we are planning a large scale study adequately controlling for all relevant variables which should give us more definitive data in regard to this important question.

Now we come to the question of the practical use of a culture-fair instrument. Critics of this type of test-instrument quite correctly point out that the measures having the best correlation with the criterion of school success are standard, or culture-bound, tests of intelligence. This is certainly a valid statement. Both our public schools and our psychological tests are biased against impoverished people. Our social structure guarantees that large numbers of poverty-area children will be unable to succeed in school and will be unable to perform adequately on tests of their "intellectual ability." There is an additional problem, however, which these critics do not mention. In typical clinical practice, especially within school systems, individual psychometric tests are administered <u>not to predict school performance but to explain school failure</u>. Hence, the nation's school psychologists spend most of their time administering intelligence tests to children who have already done poorly in school. The finding of an I.Q. of 75, for example, is used to "explain" the child's poor school performance. Condensing the school psychologists' complicated reports to simple language results in the following typical conclusion: <u>the child does not do well in school because he is limited intellectually</u>. Some of us, however, are left with the disturbing thought that perhaps many of these youngsters fail to do well in school because the school has failed to teach them. A culture-fair instrument might enable one to identify the child whose school failure is really society's failure and not his own; the lack of ability being in terms of the local Department of Education and not in terms of the child's cerebral cortex.

From this line of reasoning, I would suggest that culture-fair instruments have practical value in the following areas:

(1) Diagnosis and planning in regard to children who "fail to learn". The evaluation of intellectual potential is an important part of any diagnostic study regardless of the type of learning problem we are dealing with. Planning for such children requires one to be able to distinguish between learning difficulties due to intellectual retardation and learning difficulties due to other factors in children of normal intellectual ability. This distinction is extremely difficult to make with low-SES youngsters due to our present difficulties with standard tests of intelligence. A culture-fair instrument would be a valued addition to an evaluative battery in terms of just this diagnostic distinction.

(2) Pre-school enrichment programs, such as those already developed and those being developed for anti-poverty programs, require some means of identifying youngsters of different levels of intellectual potential. An educational program geared for youngsters of at least average potential might be expected to be totally ineffective, if not actually harmful, to retarded children. A culture-fair test would be of value in just such a screening procedure especially since we know that many current brief screening tools greatly exaggerate the number of retarded youngsters in the poverty population (e.g.

Rosenberg and Stroud, 1966).

(3) This type of intelligence measure can also be of value in educational research projects such as attempts to compare the effectiveness of different curriculum in poverty-area elementary schools. Such studies require that one be able to control for variations in the intellectual potential of the children used in the study. A culture-fair instrument can play such a role.

The personal prejudice of the author has certainly been amply expressed in the last few paragraphs. At this point, the author's bias will be made more specific. It is my belief that the intellectual ability of the poor is distributed normally. I would expect a slight rise at the lowest end of the scale due to a greater-than-expected incidence of organically based retardation related to the greater health hazards faced by the poor than by the rest of the population. In this category would be included a higher incidence of birth injury, a higher incidence of congenital difficulties related to disease in the mother, and a higher incidence of cerebral damage in young children with such causative factors as accidental head injury, lead poisoning, etc. Subtracting out this group we are left with a large number of poverty-area people whose intellectual potential distributes normally but for whom standard tests of intelligence yield I.Q. distributions which show an extreme "pile-up" in the "Borderline" and "Dull Normal" categories. This skewness of the distribution is related to the inability of current tests of intelligence to adequately measure their ability and not to an inherited inferiority of the poor.

A similar statement can be made when the data changes from I.Q. scores to levels of academic achievement. The achievement data is greatly influenced by the academic community's failure to break through the multiple barriers of social-class difference;

barriers of communication style, the lack of relevance of middle class patterns of social reinforcement, and the devastating influence of an atmosphere of apathy and despair.

Our skepticism regarding the validity of current intelligence testing results from an experience of conflict of data; too many "low I.Q." children in the lowest socioeconomic status function at a level of behavioral complexity beyond that which would be normally associated with their obtained scores. In addition, I.Q. scores of deprived-area children have shown a tendency to increase significantly when those youngsters have been exposed to even mediocre educational experiences.

In support of these statements we can briefly mention data from two sources. First, Arthur Jensen's work at the University of California at Berkeley where a learning task has been developed which has been used with children of different socio-economic class levels (Jensen, Collins, and Vreeland, 1962). When the sample is restricted to middle class children, the performance on the learning task correlates quite well with the I.Q.'s obtained from standard tests of intelligence. High I.Q. middle-class children do well on the learning task while low I.Q. middle-class children do poorly. When dealing with low socioeconomic class children, on the other hand, youngsters are found who have low I.Q.'s yet who perform at a normal rate on the learning task (Jensen, 1967). Jensen's data on middle-class children indicates that learning-rate on their instrument is related to intelligence. Hence, it is quite reasonable to suspect that the poverty-area children with <u>low</u> I.Q. and <u>high</u> learning-rate are children of adequate intelligence who have been inaccurately labeled with a "low" I.Q. score. We would predict that the disadvantaged children who performed at a normal rate on Jensen's learning task would also score high on a culture-fair intelligence test.

Related to the above findings, is data from a completely different area of investigation. Eisenberg and Conners (1966) reported that a 6-week Head Start program could produce statistically significant increases in I.Q. as measured by the Peabody Picture Vocabulary Test, Columbia Mental Maturity Scale, and Stanford-Binet. It might well be that what the Head Start study was dealing with were disadvantaged children with good learning capacity who scored very low on standard intelligence tests. Exposing these youngsters to the Head Start program better enabled them to demonstrate their actual level of intellectual ability. Hence, the post-testing shows significant increases. It appears that the laboratory work of Jensen and the field experience of Head Start studies complement each other quite well. Their data raises serious question as to the true meaning of a low score on one of our standard intelligence tests received by a disadvantaged child.

Our current problem in regard to the meaning of the I.Q. with disadvantaged children might well be aided by our consideration of a statement by Binet in 1909: " . . . . some recent philosophers appear to have given their consent to the deplorable verdict that the intelligence of the individual is a fixed quantity . . . . we must protest and act against this brutal pessimism . . . . a child's mind is like a field for which an expert farmer has advised a change in the method of cultivation, with the results that in place of desert land, we now have a harvest. It is in this particular sense, the one which is significant, that we say that the intelligence of children may be increased. One increases that which constitutes the intelligence of the school child; namely, the capacity to learn, to improve with instruction . . ."

One might equate Head Start improvement with good farming. The possible loss of the gains achieved by Head Start when these youngsters are placed in a regular school program might be equated with a situation where well nourished young plants are transplanted into very poor soil. A culture-fair instrument might well be seen as a tool which helps the farmer plan appropriately.

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.

Binet, A. cited in Stoddard, G.D.: The I.Q.: Its ups and downs. Educ. Rec., 1939, 20, 44-57; and cited in Eisenberg, L.. and Conners, C.K. (1966).

Eisenberg, L., and Conners, C.K. The effect of head start on developmental processes. Presented at Joseph P. Kennedy, Jr. Foundation Scientific Symposium on Mental Retardation, Boston, April, 1966.

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.

Jensen, A.R. Reply to inquiries concerning the article on early learning. mimeo, 1967.

Jensen, A.R., Collins, C.C. and Vreeland, R.W. A multiple S-R apparatus for human learning. Amer. J. Psychol., 1962, 75, 470-476.

LaBerge, D.L. and Lawrence, D.H. A method of generating visual forms of graded similarity. Amer. Psychologist, 1955, 10, 401.

Rosenberg, L.A., and Stroud, M. The limitations of brief intelligence testing with young children. Psychol. Rep., 1966, 19, 721-722.