A COMPARISON OF INDIVIDUAL AND GROUP ADMINISTRATION OF THE CATTELL CULTURE FAIR INTELLIGENCE SCALE AND THE WESCHLER INTELLIGENCE SCALE FOR CHILDREN IN A DELINQUENT POPULATION

by

John Frederick Ireland

A Thesis Submitted to the Faculty of the DEPARTMENT OF PSYCHOLOGY In Partial Fulfillment of the Requirements For the Degree of MASTER OF ARTS In the Graduate College THE UNIVERSITY OF ARIZONA 1974
STATEMENT BY AUTHOR

This thesis has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: John J. Ireland

APPROVAL BY THESIS DIRECTOR

This thesis has been approved on the date shown below:

M. W. KAHN
Professor of Psychology

Date 11-20-74
ACKNOWLEDGMENTS

The author wishes to express appreciation to Dr. Marvin W. Kahn, supervisor of this study, and to Drs. Phillip Balch and Alice Chang for their advice and encouragement. Thanks are due to Dr. Cecil Rogers who participated in the early planning of the study. The author is especially indebted to Mr. Greg Berg and Mr. Sam Albert who administered many of the tests. The study would not have been possible without the permission of Mr. John Kohl, Director of the Arizona Youth Center, and without the efforts of the boys who took the tests.

Finally, my wife, Andree, has not only contributed many hours of typing and editing of revisions of this thesis, but has also been a wellspring of encouragement and support and this is fully appreciated.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>v</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>vi</td>
</tr>
<tr>
<td>1. INTRODUCTION AND BACKGROUND</td>
<td>1</td>
</tr>
<tr>
<td>The Area of Interest</td>
<td>1</td>
</tr>
<tr>
<td>Two Schools of Test Development and Two Approaches to the Cultural Deprivation Problem</td>
<td>4</td>
</tr>
<tr>
<td>Intelligence of Delinquents</td>
<td>19</td>
</tr>
<tr>
<td>2. GENERAL PURPOSES AND HYPOTHESES</td>
<td>23</td>
</tr>
<tr>
<td>3. METHOD</td>
<td>25</td>
</tr>
<tr>
<td>Subjects</td>
<td>25</td>
</tr>
<tr>
<td>Procedure</td>
<td>25</td>
</tr>
<tr>
<td>File Data Study</td>
<td>25</td>
</tr>
<tr>
<td>Individual Administration</td>
<td>26</td>
</tr>
<tr>
<td>4. RESULTS</td>
<td>28</td>
</tr>
<tr>
<td>5. DISCUSSION</td>
<td>36</td>
</tr>
<tr>
<td>APPENDIX: ETHICS AND EVIDENCE</td>
<td>41</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>44</td>
</tr>
</tbody>
</table>
**LIST OF TABLES**

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Partial and Total Mean IQs and Correlations of the Cattell Culture Fair, Scale 2, and the Wechsler Intelligence Scale for Children by Administrative Conditions</td>
<td>29</td>
</tr>
<tr>
<td>2.</td>
<td>IQ Scores for File Data Study Including WISC Minus Culture Fair Difference Scores</td>
<td>33</td>
</tr>
<tr>
<td>3.</td>
<td>IQ Scores for Individual and Group Administration Study Including WISC Minus Culture Fair Difference Scores</td>
<td>34</td>
</tr>
</tbody>
</table>
ABSTRACT

IQ scores from the administration of a traditional intelligence test, the Wechsler Intelligence Scale for Children, and a "culture fair" test, the Cattell Culture Fair, Scale 2, were collected to assess the hypothesis that the mean of the Culture Fair IQ scores is higher than the mean of the traditionally determined IQ scores in a selected disadvantaged population. Individual and group administrations of the Cattell were also investigated as a source of score variance.

Ninety-six ethnically mixed, lower socioeconomic status, adolescent male juvenile delinquents served as subjects. IQ scores for forty-eight who had been given the WISC and Cattell were taken from the files of a state correctional facility. The higher WISC mean IQs in that sample led to the hypothesis that the group administration of the Cattell, compared to the individual WISC format, lowered the mean Cattell IQ scores. Therefore, forty-eight subjects entering the same facility were given the WISC and, of those, half were given the Cattell individually and half in groups of three or four.

Over all conditions, the WISC mean was higher than the Cattell mean thus placing the uniform "culture fairness"
of the Cattell Culture Fair in serious question. Further research including other variables and populations is needed to define the limits of this assessment instrument.
CHAPTER 1

INTRODUCTION AND BACKGROUND

The Area of Interest

A serious loss in human resources occurs when a society is unable to identify those with real, but untutored ability. The personal damage is also considerable as school placement, job opportunity, and the extent of clinical intervention is partially determined by intelligence estimates. Several prominent scholars feel that intelligence tests, which are used most often to make an assessment of cognitive capacity, are unfair to children from certain backgrounds which are out of the mainstream of society (Hunt, 1972; Dobzhansky, 1973; and Williams, 1974). Some children do not experience many of the situations and develop many of the skills which test constructors assume are common to all. Their performance, therefore, is lowered as a result of lack of experience, not lack of capacity.

In response to this problem, psychologists have taken two positions. One is reflected by the tradition of Binet while the other is represented by the followers of Pearson (Hunt, 1972). Pearson, as early as 1908 tried to empirically support Galton's theory of genetically distinct
and familially transmitted abilities. The modern champion of this position is Jensen (1973) who feels the poor are genetically ill equipped rather than ill trained. The factor analysts have attempted to isolate specific, as well as general, abilities for years and, when the disadvantaged do poorly on their tests, they are said to be missing a specific ability or factor. R. B. Cattell has followed this line of reasoning in developing his Culture Fair Scales (Cattell and Cattell, 1960, and Cattell, 1971).

The Binet position, on the other hand, is based upon a plastic model of intelligence holding that there are many facets to intellectual behavior. Binet, from the outset, stated that test results were gross approximations confounded with environmental opportunity. Cultural experiences and contingencies are important. Instruments developed in this tradition contain culture relevant tasks and include the Wechsler Intelligence Scale for Children (Wechsler, 1958). As opposed to the faith in factor structure which Cattell uses as a measure of test validity, the WISC relies on significant correlations with societal station and progress. If skills which contribute to a good IQ score are much the same as those required for gaining socioeconomic status, the disadvantaged, by definition, are going to do poorly. Cattell uses a dual model of intelligence where specific cultural habits and skills are "crystallized" through practice and reinforcement. "Fluid" intelligence
is biologically based and is said to be expressed in novel situations requiring original adaptive responses. Cattell contends that his test taps these physiological processes, thus bypassing the culturally mediated crystalline portion of intelligence. Latent potential is assessed and cultural deprivation effects are essentially eliminated. Cattell (1971) has explained these varieties of intelligence in a number of ways in his text:

1. "... fluid ability is conceived as a power which is a function of the total, effective, associative, cortical cell mass and of certain parameters of efficiency in those cells (p. 189)."

2. Fluid intelligence "... obviously depends primarily on sheer magnitude of neural fibre mass (p. 211)."

3. Crystallized intelligence is affected "by a different amount of educational expenditure, socio-educational status, or length of schooling (p. 276)."

4. The amount of variance due to environmental variables will "most likely be" 5 per cent for fluid intelligence and 20-25 per cent for crystallized intelligence (p. 187).

5. Genetic influences are estimated to contribute 82-87 per cent of fluid ability (p. 262).

With all of these findings it could be claimed that cultural deprivation effects are eliminated by the Cattell Culture Fair and should no longer be free to lower the
scores. Therefore, the Cattell should result in higher scores than the traditional WISC in a disadvantaged population.

Two Schools of Test Development and Two Approaches to the Cultural Deprivation Problem

The differing ideas of Pearson and Binet still exist in controversies over intelligence measurement today. The factor analytic school treats intelligence as a faculty or ability that may be broken into factors (Thurstone, 1948). Guilford has arrived at one hundred twenty abilities while others retain notions more akin to Spearman's unitary "g" (Matarazzo, 1972). It seemed to Spearman and others that the reason a group of mental tests administered to the same population showed a positive intercorrelation was because they all possessed a general factor called intelligence. This "g" is partly involved in all cognitive acts. Since perfect intercorrelations are never achieved, each task requires other more specific abilities to be fully defined. These more specific skills Spearman called the specifics or "s"s. High correlations between tasks mean they are both highly loaded on "g" and have it in common. As more "s"s become involved, the correlations lower. Spearman, like Guilford, operated in the factor analytic framework where a priori models of the nature of intelligence are used. Binet, conversely, explicitly felt trial and error was
necessary for the establishment of a valid assessment in-
strument (Matarazzo, 1972).

Binet's death allowed investigators to somewhat
erode this antitheoretical stance by introduction of the in-
telligence quotient, IQ. This index was given status as a
fixed entity by Goddard (Tuddenham, 1963), but Terman and
Merrill (1960), in their revision of the Stanford-Binet,
clearly maintain the IQ is an estimate of intellectual func-
tioning, not a final statement. The empirical school holds
that too many psycho-social variables intertwine to allow for
a fixed, absolute measure of intelligence (Matarazzo, 1972).

The issue of the cultural dependency, raised earlier
by Binet, did not recede as an issue, as descriptive studies
of the correlation of IQ with social advantage were plenti-
ful. Anastasi (1965) provides three examples. She cites
Arlitt who in 1921 tested 343 children from the primary
grades of an unidentified school district. Most, 191, were
native whites. Also included were 71 Blacks and 87 first
generation Italians. All of the latter were proficient in
English. The students were placed in five social status
categories according to parents' occupational indices. They
were professional, semi-professional and higher business,
skilled, semi-skilled, and unskilled. Over 40 per cent of
those students in the lower two social classes had IQs below
90 and 4.6 per cent had IQs above 110. Seventy per cent of
the children of the highest social class were above 110.
Very few children of the lower two status groups were of superior mentality and only one child of inferior mentality was found in the upper two status groups. Though the IQs of the native whites were an average twenty points higher than the Italians and Blacks, equating for social status reduced the difference to eight points. The Italian and Black means were lowered by a large number of scores below 80. The Italians and Blacks actually had more children of superior mentality, when equated for social class, than the native whites. In general, Arlitt concluded that social status is more important than race in the determination of IQ.

A second example cited by Anastasi (1965) is the study in 1932 by Sherman and Key who reported intelligence test scores as part of a larger project studying four mountain hollows, one hundred miles west of Washington, D. C. The hollows were close to each other, but separated by high mountains. All four of these villages differed in the degree of socialization. The lowest on the scale, Colvin, consisted of log huts and contained three literate inhabitants. Intermarriage was common. Rigby Hollow was the most developed town and was fairly easily accessible. Regular school and church activities were held and most of the inhabitants were literate. The scores of intelligence tests of children in these hollows was compared to those obtained by the children of Briarsville, a modern town easily accessible by an improved road. Briarsville had been the
recipient of several migrations from the hollow towns and had accommodated and assimilated these people. The comparison of scores was, therefore, more important as the people were of similar ancestry. The Briarsville children were higher on every test. Children of the hollow showed a decrease in IQ with chronological age and the mountain communities had a larger percentage of cases below average. Finally, the young children of the various hollows were similar in intelligence while the older children were quite different. The more socialized the hollow, the higher the older child scored. Sherman and Key concluded "that children develop only as the environment demands development."

In contrast to the cross sectional approach of Sherman and Key, Anastasi cites Wheeler who in 1942 used a longitudinal approach to study the intelligence of East Tennessee Mountain children. In 1930, he had collected data on tests administered to 1147 children and found a decrease in IQ with chronological age. The study was repeated in 1940 with 3252 children. The median IQ increased 10 points and 74 per cent of the 1930 scores were below the 1940 median. The decrease of IQ with age was again found. A great number of social innovations occurred in the decade including hot lunches, new schools, more texts and better prepared teachers. Farming methods improved and many families were able to supplement their income with jobs at new
factories. The quality of the social milieu rose as did IQ scores.

A more definitive and quite epic effort to study "Intelligence and Cultural Differences" was conducted by Eells et al. (1951). In recent years, this book has been cited minimally, but it deserves more attention. Eells et al. gave group intelligence instruments to nearly five thousand white pupils of ages nine, ten, thirteen and fourteen. Nine tests were used and all were commonly employed in schools at the time of the study. Nine and ten year olds were given the Otis Alpha Verbal, the Henmon Nelson, and the Kuhlmann-Anderson on consecutive days. The thirteen and fourteen year olds were given the Thurstone Spatial and the Thurstone Reasoning, the Otis Beta, the Terman-McNemar and the California Test of Mental Maturity on consecutive days. The tests were all given under standard conditions. At the end of the allotted time, each child was allowed to finish uncompleted questions with a colored pencil so that all test items might be analyzed. All testing was conducted in January, 1946. The pupils were essentially the entire population of that age in the Rockford, Illinois school district. Social status was determined by a questionnaire based on Warner's Index of Status Characteristics. Occupation, source of income, type of house lived in, and dwelling area in the community were weighted and combined. As cited by Eells et al., this index gave a correlation of .97 with
community observation and interview results in Warner's test sample of 209 families. Eells et al. substituted education for source of income as they felt the results would be more reliable for an impersonal mailed form. This "Rockford Index" gave essentially the same correlational results as Warner's on a 100 pupil Rockford sample. Post hoc, three status levels were derived from the data. An ethnicity tripartite was also established including "Old American," "Scandinavian" and "ethnics" categories.

Statistics computed by Eells et al. include both general correlational measures of IQ and social status and specific item analysis. At the molar level, they determined that ethnicity was unimportant compared to social class as a variable. Other conclusions included the following:

1. Correlations between IQ and status level ranged from .20 to .43. They were more highly correlated in the lower classes, less so in the higher classes. High status pupils scored eight to twenty-three IQ points higher than low status children. There was some overlap in the distributions.

2. One-half of test items for nine and ten year olds and 85 per cent of items for thirteen and fourteen year olds show status differences.

3. "Mean status differences are largest for verbal, and smallest for picture, geometric design and stylized
drawing items (Eells et al., 1951, p. 54)." Differences in types of questions showed no clear trends.

4. The largest status differences involved "bookish" vocabulary items. Simple "everyday" words showed no differences.

5. In examining patterns of errors, "a number of the differences can be accounted for in terms of status differences in opportunity for familiarity with certain objects, words, or processes.

Eells et al. (1951) conclude that though many factors intertwine, "familiarity" seemed to most adequately explain the findings. They feel their findings are important as test development rests upon the assumptions of the test constructor and the research findings to which he subscribes. If he believes genetic processes dominate intelligence and that mental intelligence structures are measurable, then he will use items with high status differences. These items would reflect the innate superiority of the upper classes and demonstrate that the lower classes are missing part of the total structure of intellect. For those choosing the developmental approach, they may take the common culture or own-culture strategies. The latter involves judging children's intelligence "in terms of the kinds of problems and experiences which are important in his own culture . . ." Since defining a cultural group seems improbable to Eells et al., they prefer an attempt at the
common culture method. One possibility they consider is giving a battery of tests, each measuring some ability in some status area. The results would then be profiled, highlighting strengths and weaknesses.

A successful test program developed along these lines has been that of Wechsler (1958, and Matarazzo, 1972). Wechsler defined intelligence "... as a hypothetical construct, which is the aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal effectively with his environment (Matarazzo, 1972)." This position is reflected in the wide variety of items within his tests and the admonition to consider all other forms of evidence of behavioral competency in arriving at an intelligence estimate (Matarazzo, 1972). In a recent monograph reviewing research with the Wechsler Intelligence Scale for Children, Zimmerman and Woo-Sam (1972, p. 37) conclude:

... the review indicates wide acceptance of the WISC as a reliable and valid test of intelligence. As is known, there is no discussion of validity in the WISC manual (1949), and this factor had been considered one of the scale's weaknesses. Nevertheless, the number of current studies that use the WISC as the validating criterion is quite large, and tests subjected to such validation include the revised Stanford-Binet L-M.

Perhaps associated with the interest in the so-called Rosenthal effect in the sixties, the roles of the examiner, examinee, and situational factors upon obtained test scores have received attention. In general, such variables do contribute to obtained scores.
Further,

Culturally deprived groups generally obtain below-average WISC scores, and thus although Negro and Mexican American children obtain mean scores below that of the standardization group, it is at this point difficult to separate the confounding effects of sociocultural variables from race. It can be argued that the WISC is not a suitable measure of intelligence for the underprivileged, but then no other standardized test (including the so-called 'culture free') is. Further, obtained scores on the WISC show substantial correlation with school achievement (p. 37).

The difficulty with sociocultural and racial variables is not particularly damaging to an instrument that has its score highly related to correlations with psychosocial competencies. This relationship is exactly what Eells et al. had hoped to avoid. Their idea of strengths and weaknesses, which they tied to cultural experience has become abortively associated with a causal genetic hypothesis. In response to this position, several authors have demonstrated that the WISC, Binet and like tests expressly measure intelligence through "WASP" activities, thus resulting in lower scores. The content "s"s, not "g" are being measured (Dobzhansky, 1973; Hirsch, 1972; Scarr Salapatek, 1971a, 1971b; and Hunt, 1972).

This position is heartily opposed by other theorists who claim that environmental variation and situational variables, when controlled, reduce minority deficits only about one-third (Jensen, 1973; Shockley, 1971; Herrnstein, 1971; and Eysenck, 1971). They conclude poor people score
lower because of innate ability deficits. The heat generated has often exceeded the light as, ironically, the chief proponent of genetic domination of intelligence variance is a psychologist (Jensen) while a leading moderate is a geneticist (Dobzhansky). The controversy will not be easily solved, as methodological questions are at base; i.e., Jensen (p. 184) feels Scarr-Salapatek's reported means are crucial while Dobzhansky (p. 21) stresses the reported variances. Interestingly, Matarazzo (1972, p. 299) reports Jensen stressing the importance of correlations in a different study while Hunt emphasized mean differences! Apparently, if a methodology produces favorable results, it is appropriate and if the results are negative the method must be lacking.

Cattell (1971), in the factor analytic tradition, has developed a theory and test of intelligence which he feels grossly minimizes cultural and situational confounding. He feels that his Culture Fair Intelligence Scales (Cattell and Cattell, 1960) tap "fluid," as opposed to "crystallized" intelligence. Fluid intelligence is operative where complex relations are involved and no prior learning is of advantage. It is highly hereditary and physiologically based. The environment cannot alter fluid intelligence and it does not increase after early adolescence. By contrast, crystallized intelligence is derived from "investment" of fluid potentialities in concrete tasks.
Crystallized abilities are the result of school and environmental experiences and may increase throughout life (Baltese and Schaie, 1974).

Cattell (1971) contends that traditional assessment instruments, such as the WISC, confound crystallized and fluid factors which result in an inability to ferret the environmental from the genetic. He stresses that the "pure" nature of the Culture Fair will enable identification of the more intelligent so that they might be provided with appropriate experiences. He also stresses that the less endowed should be afforded a "more kindly" (p. 501) curriculum. The literature of the Institute for Personality and Ability Testing, which publishes the tests, clearly does not deal with those who might not do well on the Culture Fair. One heading reads, in large bold faced type, THE PROBLEM OF HIDDEN INTELLECTUAL TALENT (IPAT, 1970). In the body of the literature, one finds parenthetical notice that the test might correct for "spurious" overestimates of intelligence; i.e., lower scores might result. This is considered as an ethical issue in the Appendix (see page 41).

The Cattell Culture Fair Scales were devised in the 1930's and have undergone numerous revisions. Scale 2, for instance, is a paper and pencil test of intelligence which is accompanied by verbal instruction in the language of the testee. The four sections of Scale 2 are typically timed and the entire administration takes approximately twenty
minutes. Series, Classification, Matrices, and Conditions are the subtests which Cattell feels comprise an adequate measure of fluid intelligence. All involve observation of geometric figures. Though a spatial ability factor may operate, its contribution seems minimal. As mentioned above, Cattell believes that for such relatively complex tasks not only are cultural and situational variables reduced in import, but that motivation will not strongly effect test scores (Cattell, 1971 and Jensen, 1973). Cattell would not seem to be distinguishing between the construct of fluid ability and the realities of test administration and test taking.

Criticisms of the Culture Fair Scales have been levied both in terms of lack of statistical data and poor conceptual reasoning. Scale 2 was standardized on 4328 subjects in Great Britain and the United States. Studies in reliability (.84), validity (.87) and alternate form equivalence (.80) are from a small sample of 102 female job corps applicants (Krug, 1967). Concurrent and predictive validity studies are sparse and range from .23 to .46 for school grades (Robb, Bernardoni and Johnson, 1972). High correlations have been found with the Abstract Reasoning section of the Differential Aptitude Battery (.78) and with the full scale WISC (.72). Cattell, however, relies on the high factor loading of "g" (.73 - .83) as a measure of the value of the instrument (Krug, 1967). Cross cultural scores have
been consistent in some cases and divergent in others (Tannenbaum, 1965). Perhaps, the most damning criticism of the statistical support of the IPAT Scales has been by Milholland (1965). He was unable to locate several references in the manuals and the ones he did find did not contain the information cited.

A related difficulty, the nonavailability of sources, appears in Barton's (1973) recent IPAT information bulletin. In it, IPAT has published mean scores of groups who took the Culture Fair as well as other intelligence tests. Three studies were cited in an attempt to demonstrate the efficacy of the Cattell. Of Barton's citations, only the study by Willard (1968) was accessible. It clearly suffers as supporting evidence for the appropriateness of using the Cattell with the disadvantaged. First, 89 Black sixth graders in New Haven, Connecticut, were administered the Academic Promise Test and the Cattell. The Academic Promise Test scores were converted to WISC IQs. The mean IQ of the Cattell was 3.6 points higher than the mean IQ of the WISC (94.4 vs. 90.8). For 42 per cent, the WISC resulted in a higher score. The range of scores for the Cattell was from 47 to 130 and the WISC range was from 71 to 124. The correlation between the two tests was .55. The reported IQs in this portion of Willard's study may be inaccurate in either direction. The WISC was not actually administered and the procedure for conversion of Academic Promise Test scores to WISC IQs was
not stated. The Cattell was administered in May, 1966, and the Academic Promise Test was given in November, 1965. The children had six months school experience between administrations and adjustment for this is not mentioned in Willard's text. The low score of 47 on the Cattell compared to the WISC low score of 71 should alert clinicians to the possibly penalizing nature of the Cattell for individual cases.

In the second part of the study, Willard tested 83 Black children in special education classes in New Haven. They were administered the Cattell and the Stanford-Binet. Information about sequence or dates of administration are not provided. The Cattell mean IQ was 1.9 points higher than the Stanford-Binet and 40 per cent of the children scored higher on the latter instrument. Willard concluded, "It would appear that the estimates of IQ determined from the Academic Promise Test total scores do not show the group in a particularly unfavorable light in comparison with the IQs obtained from the CCFI (p. 32)." Also, "The bright Negro child does well on either . . . whereas the slow or less well endowed Negro child does poorly on either type (p. 35)." Since the other two studies cited by Barton were unavailable, their populations and procedures cannot be reported. Barton cites their results which seem congruent with those of Willard. Advantages of 2.5 and 4 points were found in two samples of disadvantaged children compared to
scores on the California Test of Mental Maturity and the WISC, respectively.

The essential conceptual critique has been voiced by a number of authors in Cronbach and Drenth's book, *Mental Tests and Cultural Adaptation* (1972). They state that behavior cannot be measured independent of culture and that at least culture biased tests indicate if a person will fit into a given culture. Basic thinking processes can vary even over a few miles. At an extreme, such people as the Kalahari bushmen cannot be expected to see lines on paper as representative of meaningful material, yet the same assumptive errors may occur within a single culture context. Ord (1970) states that even the manipulation of paper and pencil is an acquired skill. Though this demonstrates the improbability of a culture fair test ever being devised, Tannenbaum (1965, p. 723) points out the attempt to control for "cultural bias in the IPAT is only a derivative of an attempt to assess latent potential rather than learned skills." The assumption which is crucial to this attempt is that nature and nurture do not interact geometrically, but are simply additive. Tannenbaum doubts that heredity could not significantly interact with environmental realities and vice versa. Early impoverishment certainly has long term effects which should be reflected in any intellectual assessment (Vernon, 1969, and Dobzhansky, 1973).
Cattell (1963), however, contends that his test measures fluid intelligence which is mostly impervious to environmental influences. The disadvantaged should not have cultural factors working against them and their scores should represent relatively pure measure of potential. More important than the theoretical state of the score is the practically and socially relevant question, "Will the score be higher?"

**Intelligence of Delinquents**

Delinquents have been cited as examples of defective intellect. Lombroso saw criminals as atavistic, reversals, a primitive life form (Rosenquist and Megargee, 1969). He suggested that a comparison of physiques would confirm this hypothesis as criminals would have such defining characteristics as low sloping foreheads. Goddard, using his own translation of the Binet scales with juveniles in a correctional facility, concluded that, "It is no longer to be denied that the greatest single cause of delinquency and crime is low grade mentality . . ." (Goddard, 1920, in Sutherland and Cressey, 1966, p. 162). Research has indicated that although an eight point IQ difference exists between delinquents and controls, confounding of social class and performance variables make any conclusions tentative (Caplan, 1965). Cattell (1971), however, has recently cited his own earlier work (1937, 1938) which posits that
it is possible that cultural complexities become "bewildering" and frustrating to the subnormal. This is essentially reflected in one dominant sociological view of the delinquent, that the unskilled delinquent seeks a subculture where rewards may be obtained and finds criminal means to that end (Cohen, 1955).

More empirically, the dominant clinical observation has been that delinquents score higher on the Performance as compared to the Verbal part of the Wechsler scales. Though group means in many studies have supported this difference, it is not appropriate for individual diagnosis (Matarazzo, 1972). A condition perhaps operating to produce the differences is a reading disability (Henning and Levy, 1967) which may result from poor impulse control (Kahn, 1968), but this is yet unproven.

In a well conceived cross cultural study by Rosenquist and Megargee (1969), several differences were noted between delinquent and control teenagers on the WISC. They gathered an extensive sociological, medical and educational data on one hundred Anglos, one hundred Latins, and one hundred Mexicans. The former groups were from San Antonio, Texas, and the latter was from Monterrey, Mexico. All were matched for social class. Extensive figures for each WISC subtest are provided by the authors. For the purposes of the present study, the Anglo and Latin samples are most important. The Anglo delinquents scored significantly
lower on Verbal and Full Scale IQs than Anglo non-delinquents. They were especially low on the information, comprehension, and similarities subtests. The Latin delinquents showed a greater disparity in comparison with their non-delinquent peers on information, arithmetic, similarities, and vocabulary. The difference was greater between non-delinquent and delinquent Latins than Anglos. This pattern was again apparent in the Performance section. Latin delinquents did significantly poorer on picture arrangement, block design, and coding. Although they did not perform well on picture arrangement, Anglo delinquents did score higher than normals on picture completion. Both delinquent groups were alike in that they did their best work on the picture completion and object assembly subtests. These findings were in general agreement with other literature (Caplan, 1965; Cropley, 1964; and Wechsler, 1958). Delinquents scored lower than non-delinquents of similar socioeconomic strata and did worse on school related subtests. The Anglo delinquents had a high average IQ while the Latins scored low average. They concluded that Latins were penalized not only for poor school competencies, but also for being out of the cultural mainstream as well.

There is only one study using the Cattell Culture Fair with delinquents. In conjunction with an evaluation program at the Arizona Youth Center, Kahn and McFarland (1973) found that after experiencing the treatment program,
mean Culture Fair scores rose from 82.6 to 88.0. Concomitantly, their reading scores and other indicators of academic improvement also rose. Thus, as with the WISC data, it is difficult to determine if the deficit (and with the other academic indicators, the subsequent rise) were due to personality, intellectual or social variables or a combination thereof. Though Cattell states the Culture Fair is a measure of fluid, innate intelligence, numerous others state that complex problem solving is linked to verbal proficiency (Lewis, 1963, and Vernon, 1969). The opposite view is also strongly represented by Jensen (1973). Jensen does indicate, in parallel to the earlier mentioned views of Cattell (1937, 1938) that the difficult content of the Culture Fair would be demoralizing to one of subnormal intelligence, thereby causing even lower test scores. The studies indicating the Cattell resulted in higher mean scores were apparently conducted on poor, but not antisocial children. In this light, only empirical investigation will decide if the Cattell is culture fair and allows the delinquent to score higher than he does on culture bound tests such as the WISC.
CHAPTER 2

GENERAL PURPOSES AND HYPOTHESES

The purpose of this research was to see if the Cattell Culture Fair Intelligence Scale would result in higher IQ scores than a conventional test, the WISC. If the Cattell corrects for cultural deficits, then scores for disadvantaged children, such as the juvenile delinquents in this study, should be higher than on a culture bound, school related test like the WISC. If the WISC scores are as high or better, then it would be hard to contend that the Cattell is culture fair. Thus, the first and main hypothesis was:

Hypothesis 1: The mean scores on the Cattell will be higher than the mean scores on the WISC in a delinquent population.

Data was available in the psychology service files and was collected to assess this hypothesis. They suggested that the higher WISC scores might be due to the individual format of the WISC which would provide incentives not operating during the usual group administration of the Cattell. After consideration of the difficult nature of the Cattell, the following hypothesis was entertained:

Hypothesis 2: When administered individually, the mean Culture Fair scores will improve and be higher than
the WISC scores. The WISC advantage due to individual administration will be eliminated.

By rejecting both hypotheses, it would establish that the Cattell, even administered under optimal conditions, is not culture fair and does not provide for higher IQ scores for one disadvantaged group, the juvenile delinquent.
CHAPTER 3

METHOD

Subjects

The subjects were ninety-six 12 to 15 year old (mean of 13 years and eight months) male juvenile delinquents who were assessed as part of the diagnostic procedures at the Arizona Youth Center psychology clinic. These delinquents came from all over the state of Arizona and a mixture of rural and urban backgrounds was represented. Racial composition included 63 per cent Caucasian, 8 per cent Negro and 29 per cent of various compounds of Mexican and Indian ancestry. A previous study indicated that each boy's family's social class level averaged between the lowest and next lowest categories on the Hollingshead scale (Kahn and McFarland, 1973). These boys had difficulty in school academically with few exceptions. Most attended sporadically at best.

Procedure

File Data Study

File data was collected by systematically reviewing the folders at the Arizona Youth Center psychology service. There were forty-eight consecutive cases selected from these
files. Records have been kept since 1970 and only those records indicating that both the WISC and Cattell had been administered were included. Though all subjects were going through the same intake processing procedure, a variety of administrative conditions was represented. The order of administration and tester consistency was indeterminate.

Means, standard deviations and correlations of the two tests were computed and compared. Table number five in the 1960 Culture Fair Manual was used to compute the IQ on the Cattell, as it was specifically designed to allow for comparisons with "contaminated" tests such as the WISC.

Individual Administration

Forty-eight consecutive juveniles who were being processed through the same diagnostic procedure as those in the file data group were randomly assigned to group or individual administrations of the Cattell. They were assessed from November, 1973, to June, 1974.

The data collected for the evaluation of individual administration was provided by giving the tests individually to one-half of the sample and in groups of three or four to the other half. Order was counterbalanced to assure both sequences appeared an equal number of times. Quasi-randomization was achieved by alternating administrative conditions from week to week. The group condition could only be administered, however, when enough clients were
scheduled. Under the individual condition, the examiner followed the instructions to the Cattell (Scale 2, Form A), but was alone with the testee in the clinic office. Under the group condition, three or four boys sat in the office at a long rectangular table about two feet apart. The examiner followed the Cattell instructions as usual. In both cases, individual boys were also administered the WISC alone in the office with the examiner. The examiner was one of three psychology graduate students on placement at the Center. The boys in the group administration of the Cattell would often not have the same examiner for the WISC, but the examiner under the individual administration condition remained constant. All the examiners were aware of the purpose of this study. Means, standard deviations, and correlations were computed and compared.
CHAPTER 4

RESULTS

The data were analyzed by a "t" test of correlated means, by correlation of the scores of each group on the two tests, and by inspection of each individual's scores. Both of the hypotheses were rejected. Under all conditions, the Cattell provided for lower scores than the WISC. Fully contrary to the second hypothesis, individual Cattell scores were significantly lower than group scores as compared to the WISC. Table 1 reflects the means of the file data (N = 48) as well as that of the data gathered under the varied administrative conditions (N = 48). Correlations between the Cattell and WISC are also provided and all are subsequently pooled and totaled.

More specifically, hypothesis 1 is clearly rejected as the Cattell mean is significantly lower than the WISC mean at the p < .001 level of probability (t = 4.28). The Cattell provided over an eleven point IQ deficit from that obtained by the WISC. As further indicated, the correlation between the two is systematically lower and significant. The Cattell had a larger standard deviation than the WISC (18.66 and 13.55), as eight scores on the Cattell were below available norms (IQ = 57). These Ss scored 75 to 89 on
Table 1. Partial and Total Mean IQs and Correlations of the Cattell Culture Fair, Scale 2, and the Wechsler Intelligence Scale for Children by Administrative Conditions.

<table>
<thead>
<tr>
<th>Source</th>
<th>N</th>
<th>CFX</th>
<th>WISCX</th>
<th>r</th>
<th>r. sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Files</td>
<td>48</td>
<td>77.58</td>
<td>89.04</td>
<td>76</td>
<td>.01</td>
</tr>
<tr>
<td>Ind. Ad.</td>
<td>24</td>
<td>84.96</td>
<td>91.79</td>
<td>83</td>
<td>.01</td>
</tr>
<tr>
<td>Grp. Ad.</td>
<td>24</td>
<td>87.75</td>
<td>91.12*</td>
<td>68</td>
<td>.01</td>
</tr>
<tr>
<td>Total</td>
<td>96</td>
<td>81.97</td>
<td>90.25</td>
<td>76</td>
<td>.01</td>
</tr>
</tbody>
</table>

*WISC was administered individually.
the WISC except for one who scored a 50. The S scoring 50 on the WISC may have been intellectually less able or malingering. One may postulate that the usual group administration condition was responsible for the large 18 to 32 point discrepancies. That is, the group situation would foster non-attention in this delinquent sample as the peers would reinforce each other for antisocial non-cooperative behavior and, thus, artificially lower Cattell scores as a function of mode of administration.

Hypothesis 2 was also clearly rejected. Under the individual administration condition, the Cattell mean was seven points below the WISC mean \( t = 3.53, p < .001 \). The difference was not as large as in the file research, but was still substantial. Tests represented in the file cases were undoubtedly administered under less homogeneous circumstances and more extraneous variables would therefore be free to operate and influence scores. Group administration provided for a closer approximation as the deficit was a little over three points. This was also a significant difference at the \( p < .05 \) level of probability. The standard deviations for the Cattell under individual and group administration were 16.48 and 12.75. WISC standard deviations were 14.79 and 12.46, respectively.

Correlations remained very high as the testees were ranked much the same by both tests. Table 1 demonstrates the correlation matrix for the two administrative
conditions. All correlations were highly significant. Thus, although the more controlled conditions reduced the deficits of the Culture Fair, the Cattell was always significantly lower than the "contaminated" WISC. Since all the differences were significant, one may pool the data for N = 96 Ss. The mean difference is then over eight points. The Culture Fair has penalized the disadvantaged, lower socioeconomic status delinquent eight points beyond whatever inequities existed in the WISC, thus, any unfairness has been compounded.

More specifically, in the file data, only ten delinquents scored higher on the Cattell than on the WISC. The mean gain was 6.8 IQ points. The 37 who scored lower on the Cattell showed a mean "loss" of 16.4 IQ points. In terms of elevating misleadingly low scores, one S had a WISC score of 74 compared to a Cattell score of 87 and one S's score of 83 was elevated to 96 by the Cattell. With respect to lowered scores, one S who had a WISC 93 scored 63 on the Cattell. One S who had a WISC 89 received a Cattell IQ of 57. In the varied administration sample, 10 delinquents also did better on the Cattell by an average of 7.8 IQ points. The 36 who scored lower on the Cattell "lost" a mean 9.2 IQ points. The Cattell IQ was 89 for one testee compared to his WISC 72. One who scored 96 on the WISC received a 71 on the Cattell. Twenty-one per cent of the delinquents, those who scored higher on the Cattell, gained an average of 7.6 IQ points.
while 76 per cent of the delinquents, those who scored higher on the WISC, lost an average 12.8 IQ points by taking the Cattell. The Culture Fair allowed an interpretive positive advantage for few. The scores of all 96 subjects appear in Tables 2 and 3.

In Tables 2 and 3 where the actual IQ scores are reported, considering those scoring below 90 on the WISC, it is evident that the Cattell aided a small number. Twenty-seven scores in the file data were below 90 on the WISC, and the Cattell elevated four of those an average 6.5 IQ points. It lowered 23 an average of 18.1 IQ points. Twenty-two scored below 90 in the varied administration groups and the Cattell benefited 9 of those an average 7.9 IQ points. It lowered 13 scores an average 10.1 IQ points. There are eight scores of 57 on the Cattell in the file data and if one eliminates them as malingerers, the losses remain an average 14.1 IQ points. All 49 scores under 90 on the WISC included, the Cattell benefited 13 an average 7.5 IQ points and cost 36 testees an average 15.2 IQ points. If those who scored under 90 on the WISC were being penalized for cultural disadvantages, the Cattell further deleted an average 9.2 IQ points. Those scoring above 90 on the WISC also lost an almost equivalent 9.0 IQ points on the Cattell.

Consistently high correlations between the Cattell and WISC also indicate that though ranking the testees similarly, the Cattell places them lower on the IQ continuum.
Table 2. IQ Scores for File Data Study Including WISC Minus Culture Fair Difference Scores.

<table>
<thead>
<tr>
<th>WISC</th>
<th>CF</th>
<th>Diff.</th>
<th>WISC</th>
<th>CF</th>
<th>Diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>78</td>
<td>69</td>
<td>9</td>
<td>93</td>
<td>87</td>
<td>6</td>
</tr>
<tr>
<td>86</td>
<td>87</td>
<td>-1</td>
<td>93</td>
<td>63</td>
<td>30</td>
</tr>
<tr>
<td>93</td>
<td>87</td>
<td>6</td>
<td>99</td>
<td>84</td>
<td>15</td>
</tr>
<tr>
<td>91</td>
<td>71</td>
<td>20</td>
<td>107</td>
<td>92</td>
<td>15</td>
</tr>
<tr>
<td>118</td>
<td>97</td>
<td>21</td>
<td>116</td>
<td>127</td>
<td>-11</td>
</tr>
<tr>
<td>77</td>
<td>62</td>
<td>15</td>
<td>72</td>
<td>63</td>
<td>9</td>
</tr>
<tr>
<td>86</td>
<td>57</td>
<td>29</td>
<td>75</td>
<td>62</td>
<td>13</td>
</tr>
<tr>
<td>78</td>
<td>59</td>
<td>19</td>
<td>88</td>
<td>57</td>
<td>31</td>
</tr>
<tr>
<td>98</td>
<td>103</td>
<td>-5</td>
<td>89</td>
<td>57</td>
<td>32</td>
</tr>
<tr>
<td>75</td>
<td>57</td>
<td>18</td>
<td>105</td>
<td>113</td>
<td>-8</td>
</tr>
<tr>
<td>95</td>
<td>97</td>
<td>-2</td>
<td>78</td>
<td>68</td>
<td>10</td>
</tr>
<tr>
<td>71</td>
<td>63</td>
<td>8</td>
<td>81</td>
<td>57</td>
<td>24</td>
</tr>
<tr>
<td>84</td>
<td>57</td>
<td>27</td>
<td>93</td>
<td>71</td>
<td>22</td>
</tr>
<tr>
<td>108</td>
<td>98</td>
<td>10</td>
<td>74</td>
<td>87</td>
<td>-13</td>
</tr>
<tr>
<td>79</td>
<td>74</td>
<td>5</td>
<td>83</td>
<td>57</td>
<td>26</td>
</tr>
<tr>
<td>86</td>
<td>70</td>
<td>16</td>
<td>85</td>
<td>77</td>
<td>8</td>
</tr>
<tr>
<td>97</td>
<td>84</td>
<td>13</td>
<td>83</td>
<td>96</td>
<td>-13</td>
</tr>
<tr>
<td>50</td>
<td>57</td>
<td>-7</td>
<td>89</td>
<td>83</td>
<td>6</td>
</tr>
<tr>
<td>77</td>
<td>65</td>
<td>12</td>
<td>83</td>
<td>62</td>
<td>21</td>
</tr>
<tr>
<td>93</td>
<td>87</td>
<td>6</td>
<td>117</td>
<td>109</td>
<td>8</td>
</tr>
<tr>
<td>90</td>
<td>92</td>
<td>-2</td>
<td>75</td>
<td>57</td>
<td>18</td>
</tr>
<tr>
<td>80</td>
<td>61</td>
<td>19</td>
<td>115</td>
<td>99</td>
<td>16</td>
</tr>
<tr>
<td>94</td>
<td>73</td>
<td>21</td>
<td>94</td>
<td>100</td>
<td>-6</td>
</tr>
<tr>
<td>109</td>
<td>109</td>
<td>0</td>
<td>88</td>
<td>60</td>
<td>28</td>
</tr>
</tbody>
</table>
Table 3. IQ Scores for Individual and Group Administration Study Including WISC Minus Culture Fair Difference Scores.

<table>
<thead>
<tr>
<th>Individual</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>WISC</td>
<td>CF</td>
</tr>
<tr>
<td>94</td>
<td>81</td>
</tr>
<tr>
<td>96</td>
<td>83</td>
</tr>
<tr>
<td>85</td>
<td>84</td>
</tr>
<tr>
<td>88</td>
<td>85</td>
</tr>
<tr>
<td>102</td>
<td>91</td>
</tr>
<tr>
<td>106</td>
<td>99</td>
</tr>
<tr>
<td>88</td>
<td>87</td>
</tr>
<tr>
<td>79</td>
<td>96</td>
</tr>
<tr>
<td>96</td>
<td>71</td>
</tr>
<tr>
<td>91</td>
<td>73</td>
</tr>
<tr>
<td>85</td>
<td>68</td>
</tr>
<tr>
<td>105</td>
<td>99</td>
</tr>
<tr>
<td>65</td>
<td>71</td>
</tr>
<tr>
<td>107</td>
<td>99</td>
</tr>
<tr>
<td>67</td>
<td>57</td>
</tr>
<tr>
<td>88</td>
<td>96</td>
</tr>
<tr>
<td>109</td>
<td>97</td>
</tr>
<tr>
<td>75</td>
<td>57</td>
</tr>
<tr>
<td>105</td>
<td>99</td>
</tr>
<tr>
<td>104</td>
<td>96</td>
</tr>
<tr>
<td>77</td>
<td>70</td>
</tr>
<tr>
<td>105</td>
<td>105</td>
</tr>
<tr>
<td>119</td>
<td>118</td>
</tr>
<tr>
<td>67</td>
<td>57</td>
</tr>
</tbody>
</table>
The WISC affords the delinquent an opportunity to display more of his abilities. This was confirmed across races in an analysis of the scores of those in the varied administration conditions. The four Blacks had mean scores of 80.5 on the Cattell and 86.25 on the WISC. The fourteen testees who were of Mexican-Indian descent scored an average 80.4 on the Cattell and 87.9 on the WISC. The thirty whites in the study scored 89.9 on the Cattell and an average 93.8 on the WISC. As demonstrated in prior discussion, those already scoring lower than average on the WISC are those most penalized by the Cattell.
CHAPTER 5

DISCUSSION

Results show that scores on the Culture Fair are systematically lower than scores on the WISC for a lower class, ethnically mixed, delinquent population. Individual administration is not a major factor in the difference in this delinquent population as was considered in hypothesis 2. It is, therefore, difficult to maintain that the Cattell is truly more fair than the traditional instrument. In fact, in this population of delinquents, these disadvantaged people actually did worse, not better, on the Culture Fair. Further, the lower scores were uniform, as the tests correlate highly. The rankings of the testees were similar, but the scores were systematically lower.

The consistent superiority of the WISC scores would indicate that the Cattell is simply a more difficult test for this group. Cattell (1971) feels that the test is rightfully hard, as those with average or better ability will have possession of the necessary skills because of inherited neuronal efficiency. Genetic engineering alone will allow permanent rises in a population's intelligence (p. 295). This is also the position of the Jensenist camp.
An opposing camp contends that certain minimal environmental experiences are needed for one to be able to think symbolically (Deutsch, 1968; Bernstein, 1960; Vernon, 1969; and Lewis, 1963). Because the Cattell was based on the former premise, any reduction in culture bias was "only a derivative of an attempt to assess latent potential rather than learned skills (Tannenbaum, 1965)." Cattell's (1971) recent book and IPAT test material do not realistically consider the possibility that the Cattell could result in lower scores. In retrospect, the evidence presented as proof of culture fairness was the equivalent scoring patterns cross-culturally. Means and standard deviations were similar in different countries. It makes sense that a very difficult test of geometric content would result in equivalent patterns. The "bright" score well and the "subnormal" score badly (Willard, 1968). The cultural variables are still at play. The concepts have just been learned in different languages. Only the utility of carrying one test instead of several could possibly be gained.

The losses to the disadvantaged are multiple. The logic of the Cattell is that one gives a hard test to be fair. As delinquents are often deficient in reading skills (Henning and Levy, 1967), they are, perhaps, also deficient in building blocks necessary for symbolic thinking (Deutsch, 1968). Concrete thinking is, in fact, more adaptive in
their environment (Bernstein, 1960). Therefore, as one moves toward the abstract, he multiplies any deficits the delinquent may have. Eells et al. (1951) demonstrated that abstract thought is exactly where lower socioeconomic groups have difficulty. They state "the middle class handles chiefly symbols for a living, while the lower class handles chiefly things. Since low status pupils are less likely to hear and be familiar with abstract words and less likely to be accustomed to thinking in abstract terms, the results . . . are not surprising (p. 284)." More current work of Hess and Shipman (1967) and Hellmuth (1970) has reached essentially the same position. Lewis (1963) has postulated that the transition from Piaget's concrete to formal operations never fully occurs in disadvantaged populations. Further, delinquents are often recommended for "structured" environments in order to develop internal sources of motivation and discipline (Redl and Wineman, 1951, and Stubblefield, 1967). The WISC uses relatively common content and the examiner remains in active contact with the testee. The examinee may be free to move about while answering verbal questions and, on the performance items, has concrete materials to manipulate. The Cattell is more ambiguous in structure as the testee is left to his own devices for the testing periods. Finally, his locus or topology of response is limited to sitting in a chair and marking answers with a pencil. The same school type
behavior which the delinquent performs poorly is called for in taking the Cattell. If actual geometric objects were used, instead of printed figures, the Cattell might be a more appropriate instrument. Pictures of the "real" objects might be mounted on slides and selected by mechanical means as a more engaging mode of presentation.

As is, one would expect the Cattell to be more difficult for the disadvantaged delinquent. The fact that the Cattell IQ was even lower than the WISC verbal IQ in this study, makes the above explanation especially compelling. The clinical implication is clearly that the Cattell Culture Fair is much more penalizing than the WISC in assessing disadvantaged delinquents.

Those who consider the Cattell for institutional use should not minimize the demand characteristics of the instrument. Under controlled conditions, it more closely approximated WISC scores, though never equalled them. The more similar mean scores in the group administration condition might have been due to copying of the smartest child's answers by others in the group unsure of their own abilities. Those who consider the Cattell because of brevity and economy will have to assess the import of possibly lower scores and the hypothesized copying phenomenon on the fortunes and futures of their clients.

Of course, this study has limitations which should also be considered. The population was ethnically mixed
and represents only institutionalized juvenile delinquents in one state. Grossly unequal Ns make generalizations about racial influences difficult. All the test administrators were white as a further confounding factor. The administrators of the tests, also, were aware of the purpose of the study and might have influenced the results. Finally, the results raise many new questions, such as what differentiated those few delinquents who did score higher on the Cattell? Do the results generalize just to delinquent populations or the disadvantaged in general? What variables accounted for Willard's (1968) more positive results? Further research is needed into these questions as well as into the search for an abbreviated, yet fair, instrument for intelligence assessment.
APPENDIX: ETHICS AND EVIDENCE

IPAT has published the Cattell measure of intelligence not as one of a genre of tests attempting to be "culture fair," but as the "Test of 'g' Culture Fair." Though it is the responsibility of psychological examiners to be aware of the limits of an assessment instrument, the publishers, as well, should be equally aware and explicit. If one names a test "Culture Fair," ample evidence should be martialed to support this contention. IPAT has not offered substantial support for the Cattell.

Since the Cattell Culture Fair was originally published in 1933, studies comparing mean IQs from its administration and those of other instruments have been few. To this author's knowledge, there is no empirical support for the presumption that Cattell administration results in significantly higher scores for any population or group. This is extremely important as IQ scores are often used, in part, to make decisions about educational goals and employment potential. The State of Arizona, for instance, enters IQ scores as part of the juvenile delinquent's permanent file. One residential placement facility for juveniles requires IQs of 100 to enter their program. The three studies reviewed by Barton (1973) and discussed previously are not
compelling in demonstrating minimally higher IQ scores from children administered the Cattell versus other instruments. Recently, Watson and Klett (1974) compared the IQ scores resulting from three tests and found the Cattell resulted in the lowest mean scores. The Cattell average IQ was 88.1, the Porteus IQ was 120.2, and the WISC average IQ was 103.6 for this sample of V.A. patients. Coupled with the results of the current study, the evidence now available actually makes the Cattell seem unfair and penalizing.

That empirical evidence has been sparse is most likely due to the logic used by the creators of the Culture Fair. The reasoning was, apparently, that if administrations in different countries yield similar distributions of scores, then the test is culture fair. If it is culture fair, then they assume it must be fair to subcultures within a larger culture and compensate for environmental disadvantage. This reasoning is fallacious, however, as the similarities across countries is based, in part, upon a similar distribution of environmental advantages and disadvantages. In both countries, then, the deprived would do poorly and the more affluent would do well. Therefore, until the data presented above was gathered, no predictions concerning the efficacy of the Cattell could be made. Cross cultural evidence is not fully relevant, if relevant at all. Paragraph headings such as "THE PROBLEM OF HIDDEN INTELLECTUAL TALENT" are inappropriate as the only relevant data,
that mentioned above, would deny the Cattell any efficacy as a detector of latent ability. In fact, data from this study would indicate that the Cattell further penalizes those already suffering from environmental disadvantage.

In consideration of the empirical evidence and the logic presented above, IPAT should consider the modification of their promotional literature and the abandonment of the label, "Culture Fair." If not, Principle 5, public statements, Principle 10, announcement of services, and Principle 15, test publication may well be violated (American Psychological Association, 1967). The label "Culture Fair" gives the test a superficial validity not formally validated by data. The brochures are less than modest and are persuasively written. This may be normative practice by the publishers of psychological assessment instruments, but it seems ethically and strategically unwise in a profession asking to take an increasing role of service, trust and responsibility in the health care of the nation.
REFERENCES


Scarr-Salapatek, S. Race, social class, and IQ. Science, 1971a, 174, 1285-95.


