The University of San Francisco

A COMPARISON OF THE BLACK INTELLIGENCE TEST OF CULTURAL HOMOGENEITY WITH THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN (REVISED), AS MEASURED BY A CONVENTIONAL ACHIEVEMENT TEST WITHIN A BLACK POPULATION AT DIFFERENT SOCIAL CLASS LEVELS.

A Dissertation

Submitted to

The Graduate Faculty in Educational Psychology

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Education

By

Richard A. Komm

San Francisco, California

April 1978

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

This dissertation, written under the direction of the candidate's dissertation committee and approved by the members of the committee, has been presented to and accepted by the Faculty of the School of Education in partial fulfillment of the requirements for the degree of Doctor of Education.

Richard A. Konan
Candidate

April 28, 1978
Date

Dissertation Committee

Chairperson

Lawrence E. Murphy

John R. Harris

Virginia B. Woodson

Arthur Fisher
Dean, Graduate Division

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ACKNOWLEDGEMENTS

This writer wishes to thank all those individuals who gave so generously of their time and effort so that this study might be completed. In addition, Dr. H. D. O'Neill, Jr., as faculty advisor, Dr. Thomas McSweeney, as Department Chairman, and Dr. James Colwell offered encouragement and advice at critical junctures during the evolution of this study.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	i
TABLE OF CONTENTS	ii-iii
LIST OF TABLES	iv-v
CHAPTER I. INTRODUCTION	1
Statement of Problem	1
Purpose of Study	9
Hypotheses	13
Overview of the Study	15
Population and Sample	16
Procedure of the Study	17
Analysis of Data	19
Limitations of Study	20
CHAPTER II. REVIEW OF RELATED LITERATURE	22
Introduction	22
Early Theoretical Development	22
The Controversial Intelligence Test: Historical, Social and Ethical Implications	25
The Problem of Minority Group Testing and Its Application to This Study	45
Criterion-Referenced Testing	52
Summary	54
CHAPTER III. DESIGN OF THE STUDY	56
Introduction	56
The Sample	56
The Instruments	63
Test Administration	64
Reported Reliability and Validity Indices	64
Standardization Sample	69
Reliability and Validity Data	71
Standardization Design and Procedure	74
Score Reportage and Concurrent Validation	74
The Method	75
Testing Procedure	78
Procedural Followup	80
Additional Information	81
Chapter Summary	82

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Page

CHAPTER IV. STATISTICAL ANALYSIS 83

Overview 83
Preliminary Considerations 83
Restatement of Hypotheses 85
Statistical Method and Rationale 85
Descriptive Statistics 88
Inferential Data. 91
Summary 93

CHAPTER V. DISCUSSION OF RESULTS 104

Introduction 104
Data Relevance and Its Relation to the Study's Hypotheses. 104
Additional Findings 108
Implications and Conclusions 110
Limitations of the Study 112
Concluding Summary 113

BIBLIOGRAPHY. 114-128

APPENDIX. 129

The Black Intelligence Test of Cultural Homogeneity

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LIST OF TABLES

Table		Page
I.	The Meier Bell SES Scale	60
II.	Residence Scale of SES	61
III.	Total SES (Social Class) Rating Scale	62
IV.	Means and Standard Deviations of Blacks and Whites on the BITCH 100	64
V.	BITCH 100: Split Half and Test-Retest Reliabilities	65
VI.	Coefficients of Correlation Between BITCH and CAT Scores	66
VII.	Verbal and Performance Scale of the WISC-R	68
VIII.	Standardization Sample by Geographic Region and Race as Compared to the 1970 U.S. Population Characteristics	70
IX.	SPSS Coding Data (Legend)	84
X.	Scores, Frequencies, Adjusted and Cumulative Percentages of all Ss on the BITCH	89
XI.	Score Differentials by Social Class on the WISC-R	90
XII.	Subscale Mean-Differentials on the WISC-R by Social Class	90
XIII.	Grade Score Differentials by Social Class on the CAT-Reading Test	91
XIV.	Means and Standard Deviations on all Correlated Variables	92
XV.	Intercorrelations of Variables and Significance Levels for All Ss (N=65)	94-97
XVI.	Kendall Correlation Coefficients and Significance Levels on All Variables Within the Study (N=65).	98-101

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table Page

XVII. Analysis of Variance Scores, F Statistic, and Significance Levels Between Social Class (By Category) and Study Variables (N=65) 102-103

XVIII. T-Test Results to Determine Significance Between Male/Female Sample Means 103

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER I

INTRODUCTION

Statement of Problem

Although the problem of cultural bias of various test instruments has always been present in test construction and validation, in recent years it has become acute, especially in relation to the identification of mental abilities in minority group children. Myriad aspects of the "nature-nurture" controversy over intelligence and its measurement have been debated since before the turn of the century with the coining of the term "mental test" by J. McKeen Cattell in 1890.[1] The controversy intensified with the introduction of the Stanford revision of the Binet Intelligence scale in 1916 at a time when mental testing had gained great notoriety. Much of the popularization of testing occurred with the development of the first large scale group intelligence tests which paralleled America's entry into World War I.[2] The Lippman-

---

[1]John Loelin, Lindzey Gardner, J. Spuhler. I. N., Race Difference in Intelligence, (San Francisco, Ca.: W. H. Freeman, 1975), p. 4.

[2]Anne Anastasi, Psychological Testing, Fourth Edition, (New York, N.Y.: Macmillan, 1976).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Terman debates provide ample evidence of the polarization of those points of view which characterize and differentiate the geneticists from the environmentalists.[3] Although this same issue has emerged repeatedly in the literature over five decades, as indicated by Cronbach in his summary article,[4] Horn states, "the perennial (if often dormant) interest in the heritability of intellectual abilities was revived by the wide dissemination of a Harvard Educational Review article by Jensen."[5] Among other aspects of Jensen's study was the reiteration of the results of Burt and others regarding the inheritance of intelligence.[6] This article also tended to underscore similar findings by Shuey, particularly as applied to Negro-White differences in intelligence over a fifty year span.[7] Her contention was that an examination of approximately 400 studies in this area indicated a fifteen point discrepancy on intelligence tests in favor of Caucasian populations. Moreover this difference, "all taken together,

---

[3]N. J. Block, Gerald Deworkin, The I. Q. Controversy: Critical Readings, (New York, N.Y.: Mac Millan, 1976).

[4]Lee J. Cronbach, "Five Decades of Public Controversy Over Mental Testing, "American Psychologist, (1975).

[5]J. L. Horn, "Human Abilities, "Annual Review of Psychology, V. 27 (1976), p. 473.

[6]Arthur Jensen, "How Much Can We Boost I. Q. and Scholastic Achievement?" Harvard Educational Review, V. 39 (1969), pp. 1-123.

[7]Audrey Shuey, The Testing of Negro Intelligence, Second Edition, (New York, N.Y.: The Social Science Press, 1966).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

inevitably points to the presence of native differences between Negroes and Whites, as determined by intelligence tests."[8] Jensen's statements that the heritability of intelligence accounted for .80 of an individual's ability, and that Black populations are lower along this dimension, touched off a controversy which is still raging.[9] The list of books and articles published which either support or denounce Jensen's conclusions are voluminous (see Bibliography). Horn indicates, "the evidence is equivocal and far from adequate to provide a convincing case for either a mainly genetical or a mainly environmental interpretation."[10]

Out of the welter of rhetoric stemming from the controversy are more specific issues focusing upon the use of ability tests which were reported to unfairly discriminate against Black populations. This was felt to be due to the presence of test and/or user bias which negatively influenced and consigned Blacks to inferior positions in schools or other environments,[11] thus perpetuating "Black intellectual

---

[8] Shuey, Ibid., p. 521.

[9] Arthur Jensen, Genetics and Education, (New York, N.Y.: Harper and Row, 1972), pp. 1-67.

[10] J. L. Horn, Op. cit., p. 473.

[11] Robert L. Williams, "Danger: Testing and Dehumanizing Black Children," Clinical Child Psychology Newsletter, V. 9, #2S (1970).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

genocide."[12] The Association of Black Psychologists and others have not only written of the negative effects of attaching I. Q. labels to Black children, but have proposed alternatives to such practices (Rivers, et al, 1975; Samuda, 1975). Williams has indicated various alternatives to "standardized intelligence tests" which, in his opinion, contain a "built-in" test bias against Black children.[13] He has developed an instrument called the "Black Intelligence Test of Cultural Homogeneity" (BITCH) which is "culture-specific" in nature, pertaining more or less exclusively to a particular sub-culture, in this case the Black population of the United States.[14] Williams cites Barnes in stating,

> "...the point is that culture specific tests could be used to determine the child's ability to function symbolically or to think in terms of his own culture and environment... If he can learn (various) relationships in his own culture, he can also master those aspects of the elementary school curriculum requiring this dimension of ability..."[15]

In discussing the BITCH-100 as a culture-specific, Williams states that the instrument is not intended to be

---

[12]Robert L. Williams, "From Dehumanization to Black Intellectual Genocide: A Rejoinder, " Clinical Child Psychology Newsletter, V. 9, #3F (1970), p. 8.

[13]Robert L. Williams, "Moderator Variables as Bias in Testing Black Children," Journal of Afro-American Issues, V. 3, #1, (1975), pp. 77-90.

[14]Robert L. Williams, "The BITCH-100: A Culture-Specific Test," Journal of Afro-American Issues, V. 3, #1, (1975), pp. 103-116.

[15]E. Barnes, "I.Q. Testing and Minority Children: Imperatives for Change," National Leadership Institute Teacher Education, (University of Connecticut: Technical Paper, 1972), p. 6.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

culturally fair and, with all items taken from the Black experience, such a test, "has the advantage of dealing with content material which is familiar to the Black child."[16] He then makes several related statements, namely that this instrument:

a. Can be used to rule out persons who have been determined as mentally retarded or low I.Q.

b. A high score on the BITCH contradicts a low score on the WISC or Binet.[17]

It appears that these remarks deserve to be studied more carefully, both from the point of their implications, as well as from differing empirical data which has been published. In the former case, Williams clearly implies that the BITCH is, in effect, a test of intelligence. In point of fact, is the BITCH actually a test of intelligence, or, rather a measure of highly specific experiential data? Williams admits that he is not exactly sure just what his test measures stating:

> "...Next we come to the knotty problem of validation. How do we know that the BITCH is measuring intelligence rather than some other phenomenon? My honest reply is that I do not know..."[18]

His second point which touches on the fact that a BITCH high score contradicts a WISC or Binet low score is contradicted

---

[16]Robert L. Williams, Op. cit., p. 108.

[17]Robert L. Williams and Associates, Cover Letter for BITCH Test Kit, (St. Louis, Mo.: Williams & Associates, 1973).

[18]Williams, Op. cit., p. 112.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

by several studies (Long & Anthony, 1974; Wright, 1975). Both of these studies showed that retarded Black students obtained similar scores on the WISC and the BITCH.

Williams eschews traditional validation procedures, indicating that standardized intelligence tests, when used as predictor variables, are biased against Black populations.[19] However, his correlations with criterion measures (achievement tests) are low, with BITCH high scorers correlating from +.18 to +.39 with the California Achievement Test (CAT). If, as has been oft-stated, standardized intelligence tests are more closely related to school aptitude (Gallagher & Moss, 1963) or adequately predict scholastic achievement (Butcher, 1968), the fact that the BITCH correlates minimally with achievement may well indicate that its apparent inability to adequately predict criterion measures seriously hampers its effectiveness as a useful discriminator of "intellectual indicators"[20] in Black populations. Certainly this important aspect must be further investigated; there exists a seeming contradiction in William's contention that ability testing must be related to its sensitivity in predicting criterion measures,[21] and his following statement:

> "...As might be expected, there is a low correlation between the three subtests of the CAT and the BITCH. Thus the Ss who scored low in the CAT did

---

[19]Williams, Op. cit., pp. 77-90.

[20]Williams, Op. cit., p. 114.

[21]Williams, Ibid., pp. 103-114.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> not necessarily score low on the BITCH. To the contrary, some of the low CAT scorers were among the high BITCH scorers. These findings suggest that the BITCH and the CAT may be measuring different phenomena."[22]

A related but important distinction that is addressed in the literature surrounding the culture-specific testing movement relates to the fact that the instruments are normed on urban Black samples using stimuli that may well be highly localized and germane to that particular region or locale only. On the basis of this localized sample, Williams then generalizes the applicability of the BITCH to the Black population of the United States.[23] Moreover in discussing the BITCH, one of its major points relates to the fact that "disadvantaged" Black populations answer culture-specific items with relative ease. This indicates that, "if the individual has the ability to learn in his native environment, he can learn in another."[24] Does this generalization also hold true for middle-class Black groupings? In the BITCH test manual, Williams reports that "half the Ss were from low socioeconomic levels, whereas the other half came from middle income levels."[25] He fails to specify that income level alone truly differentiates social class identification. Moreover in assessing his

---

[22]Robert L. Williams, Black Intelligence Test of Cultural Homogeneity, Test Manual, (St. Louis, Mo.: Williams and Associates, 1972), p. 11.

[23]Williams, Loc. cit., 1973.

[24]Williams, Loc. cit., 1973.

[25]Williams, Loc. cit., p. 6.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

data, no breakdown is given to differentiate scores between low and middle class Black samples.

The need to apply additional variables, rather than relying solely on family income in distinguishing social class has been aptly summarized by Yee.[26] He reports a study (Kahl & Davis, 1956) which found that the occupational level, education and residence (home types) of the parents are extremely significant variables which must be taken into account in differentiating social class data. In determining their indices, Kahl and Davis factor-analyzed nineteen separate stratification criteria. Similarly, Kohn (1970) "found the interaction effects between social class and income to be very small and thus insignificant. He concluded that income (taken by itself) would add very little in identifying social class position."[27] In view of the fact that Williams has apparently relied solely upon income as a measure of social class status, it is not at all clear that he has accurately discriminated social class variables among his norming population on the BITCH. Hence his statements generalizing data for Black populations at all social class levels throughout the country are called into question.

---

[26]Leland Yee, Mental Abilities of Chinese Fourth Grades at Different Social Class Levels, Unpublished Master's Thesis, (San Francisco State University, San Francisco, Ca., 1972).

[27]Yee, Ibid., p. 33.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Purpose of the Study

Taking into account the preceding information, it seems appropriate to initiate a study which will shed light on the following:

a. How well does the BITCH correlate with a well-researched and established measure of intelligence that is normed for the population in question?

b. Is the BITCH a valid predictor of scholastic achievement, (as would be the case with a standardized and established measure of mental ability)?

c. Is the BITCH an equally valid discriminator of "Black intelligence" between low and middle social class groups within the Black population?

Data pertaining to the initial statement cited above would directly relate to the concurrent validity of the BITCH, especially as it "taps" those intellectual indicators which the two instruments share. With regard to the second statement, data forthcoming on this point would further clarify the concurrent validity of the BITCH when related to the criterion measure of scholastic achievement. Information relating to the third statement would reveal variance within groups with regard to the moderator variable of social class within a racially similar populations.

It is proposed to study the above-mentioned interactions in a target sample of Black adolescents. No attempt will be made to control for examiner race or sex, since a number of studies have indicated little or no significance in matching the race of the tester and testee with relation to test performance (Jacobs & DeGraaf, 1973; Wellborn, et al,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1973). Sattler reports,

> "...When research studies are reviewed, the results suggest that the examiner's race does not usually affect the performance of Negro or white subjects on individual or group-administered intelligence tests."[28]

In viewing sex differences, it is noteworthy that Williams does not separate his sample along this dimension, although the study by Long and Anthony (1974) did uncover higher scores by females on the BITCH which were significant at the .01 level.[29] Thus a target sample divided equally between males and females would investigate reported sex differences in this area as well.

The use of the Revised Wechsler Intelligence Scale for Children is proposed as the instrument of concurrent validation for several important reasons:

a. It is a well-researched and widely accepted measure of general intelligence within its standardization population; both this instrument and its predecessor (WISC) are well-reported in the literature (Buros, 1972; Anastasi, 1976).

b. The recent revision of the WISC-R (1974) included representative samples of non-white populations according to the prevailing percentages of such individuals as reported by the most recent (1970) census figures (Wechsler, 1974).

---

[28]Jerome Sattler, Assessment of Children's Intelligence, (Philadelphia, Pa.: Saunders, 1974), p. 33.

[29]Peggie Long, John J. Anthony, "The Measurement of Mental Retardation by a Culture-Specific Test, "Psychology in the Schools, (1974), p. 312.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The latter statement is particularly relevant since one of the major objections of its detractors is that the initial version of the WISC did not include standardization samples for Black children in its original norming population (Wechsler, 1949). Since the revised WISC does indeed contain data on Blacks and other non-white populations in direct proportion to the most recent (1970) census figures for non-whites (15.0 percent), resultant figures between this instrument and the BITCH should provide relevant comparative data. Moreover, in terms of sample representation of non-whites, 92.4 percent are Black which very closely parallels and slightly exceeds the 91.3 percent representation of Blacks in the United States' non-white population.[30] This is especially crucial since Williams, in commenting on the WISC revision states:

> "Reports indicate that the Psychological Corporation is re-standardizing the WISC using 10-12 percent (or stratified sampling techniques) of Blacks and other minorities in the sample. This is clearly unacceptable. First of all, minorities comprise about 16-20 percent of the total population. A sample of 10-12 percent is not representative..."[31]

As can easily be seen from the norming data, such a statement is in error and calls into question the theoretical assumptions underlying Williams' apparent refusal to compare the BITCH with the revised WISC. Certainly such a comparison could only serve to illuminate and objectify any real or

---

[30]David Wechsler, Manual for the Wechsler Intelligence Scale for Children - Revised, (New York, N.Y.: Psychological Corporation, 1974), pp. 19-23.

[31]Williams, Loc. cit., p. 80.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

hypothesized differences between the two instruments.

At this point, the reader may well wonder why the BITCH was chosen to cross-validate with the WISC-R. Several key factors must be briefly reviewed; first, in relation to the test itself; second, in relation to its creator, Robert L. Williams. Of the entire repertory of culture-specific tests for Black populations, the instrument that receives the most attention is, by far, the BITCH. It has been studied and researched both on retarded populations (Long & Anthony, 1974; Wright, 1975) and in comparing normal Black and Caucasian samples (McNiel, 1975; Williams, 1972). It has been partially cross-validated by regions (Boston, Mass.; Mississippi State; St. Louis, Mo.; 1972). It is mentioned by Samuda as being suitable and appropriate for minorities and "educationally disadvantaged."[32] A most recent work by Oakland (1977) devotes space to a discussion of culture-specific tests, of which the BITCH is given primacy over all other such instruments, as applied to Black populations.[33]

In relation to its creator's reputation, Dr. Robert L. Williams is currently Professor of Psychology, Washington University, St. Louis, Mo., as well as being Director of Black Studies at that institution. Dr. Williams is also President

---

[32]Ronald Samuda, Psychological Testing of American Minorities, (New York, N.Y.: Harper and Row, 1975), pp. 145-180.

[33]Thomas Oakland, Psychological and Educational Assessment of Minority Children, (New York, N.Y.: Bruner/Mazel, 1977), p. 15.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of the National Association of Black Psychologists, and, as contributing editor of the Journal of Afro-American Issues, has written numerous articles condemning the use of standardized intelligence tests with Black populations (Williams, 1970; 1972; 1973; 1974; 1975). As the chief proponent of the culture-specific test movement in both the literature and by means of having created the BITCH as a prime example of his theoretical orientation, Dr. Williams' theories merit empirical investigation. By subjecting the BITCH to comparative study, the relevancy and significance of the culture-specific viewpoint can be clarified.

Hypotheses

Hypothesis I: There will be a significant relationship between the BITCH and the verbal scale of the WISC-R.

Explanation Ia: The revised WISC, unlike the earlier version, contains norms which include an appropriate Black sample. Due to this fact, the resulting standardization data represents variables common to a portion of the Black sample which will appear on both instruments.

Explanation Ib: A series of factor analytic studies performed on the original WISC yielded evidence of three primary factors which correspond to verbal comprehension, perceptual organization, and memory (freedom from distractability).[34] These

---

[34] I. Zimmerman; J. M. Woo-Sam, "Research with the Wechsler Intelligence Scale for Children, 1960-1970," Psychology in the Schools, V. 9, (1972), pp. 232-271.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

factors were similar, even when separate racial and ethnic groups (Caucasian, Black, Mexican-American) were analyzed (Silverstein, 1973). A quite recent study on the WISC-R revealed similar factors corresponding closely to those previously cited.[35] In view of this data and especially in light of the BITCH item selection, consisting of a matching word list,[36] it appears highly likely that a verbal comprehension factor is being tapped. For this reason, comparative data between the BITCH and the verbal scale of the WISC-R should yield significant correlations.

Hypotheses II: There will be a significant relationship between the BITCH and Verbal Achievement, as measured by a conventional Achievement Test.

Explanation: The concurrent validity of the BITCH with achievement measures will occur along similar lines to that anticipated in Hypothesis I. To the extent the verbal comprehension factor loads the BITCH and has bearing upon similar factors which underlie portions of relevant achievement tests, a significant correlation will be seen.

Hypothesis III: There will be a significant relationship between middle and lower Black social class groupings on the BITCH,

---

[35]A. S. Kaufman, "Factor Analysis of the WISC-R at Eleven Age Levels Between 6 1/2 and 16 1/2 Years," Journal of Consulting and Clinical Psychology, V. 43, (1975), pp. 135-147.

[36]Williams, Op. cit., p. 108.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

with low social class Blacks scoring higher than middle SC individuals in the test sample.

Explanation: This touches on the lesser degree of familiarity that middle SC Blacks have of items culled largely from terms more representative on inner-city Black urban populations. The possibility must be considered that, along this variable, the BITCH is less representative of the Black population as a whole and more germane to lower SC urban Blacks only. The question is therefore raised if the BITCH is equally valid for middle, as well as lower Black SC groups.

Overview of the Study

The purpose of this study, therefore, is basically an attempt to assess concurrent validation measures between several measures of verbal intelligence and achievement. A comparison will be made involving a relatively new culture-specific instrument expressly created for Black populations (BITCH) with the Revised Wechsler Intelligence Scale for Children (WISC-R) whose recent restandardization includes a representative Black sample. Data obtained from these instruments will then be compared to achievement (criterion) variables among Ss within a Black sample. These factors relate to the sensitivity of the instrument in determining intellectual indicators equally among middle, as well as lower social class Black subjects. In other words, it is important to ascertain whether the BITCH is actually an "intelligence test" for Blacks, as its creator claims, or whether it is actually a measure of understanding "street talk" and related terms.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Population and Sample

The population consists of Black adolescents native to the San Francisco Bay Area. It is felt that by drawing a sample from an exclusively Black population, interaction effects occurring due to the race of Ss could be eliminated. Thus the results obtained from such a sample will be a purer measure of what is inherent within the instruments themselves. The sample will consist of sixty (60) Black adolescents between the ages of 13 1/2 and 17 years. This range fulfills the norming criteria for the WISC-R, since the standardization population for this instrument is from ages 6 to 17 years. With regard to the age range on the BITCH, Williams notes that the test "is to be used primarily for adolescents and adults."[37] Of the sixty Ss, thirty will be male and thirty female.

With regard to the social class differentiation, the variables of income, occupational level, education, residential locale and type of residence of the parents will be used to separate middle from lower social class identification in the sample. Scales will be prepared delineating the above variables at the relevant indices of parental social class identity. Both the description of the scales and the rationale for this approach will be described more fully in a later portion of this study.

The sample will consist entirely of Ss currently at-

---

[37]Williams, et al., Cover Letter for the BITCH Test Kit, (1973).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

tending school in the San Francisco Bay region; the sample will also be of urban (large city) make-up. Only those Ss currently succeeding at their age-appropriate grade levels will be utilized; thus an essentially "normal" sample will be targeted for this study. The further delineation of the sample and the rationale involved will also be more fully explored in a later section of this report.

Procedure of the Study

After selection of the Ss along previously-mentioned lines, a contact will be made between both the Ss, as well as their parents. This contact will be aimed at securing the permission and cooperation of all parties involved. A statement will be prepared which must be signed by the parents; at the time of their signing the permission, the relevant data pertaining to the social class scales will also be obtained. Additionally, the families can be questioned as to their length of residence in the San Francisco Bay Area.

The Ss will be seen within the site of the school of their attendance; for this reason, the cooperation of the administrators and on-site counselors will also be secured by prior contact and discussion of the nature, as well as the importance of the study. Ss will take either the BITCH, WISC-R, or the CAT-R initially, since the instruments will be administered in randomly counterbalanced order, to control for ordering effects. Since Williams' original norming groups used the California Achievement Test, this same instrument will be used in this study. In reviewing the original data

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

on the CAT, it is noted that an r of .18 (insignificant) was obtained when BITCH scores were compared to CAT arithmetic measures.[38] Due to the highly verbal nature of the BITCH (see Chapter III), it is not anticipated that correlations of any significance will be obtained between the BITCH and the CAT arithmetic measures on this study. Accordingly, only the appropriate measures on the CAT (i.e., reading vocabulary and comprehension) will be used in this research. Although it is anticipated that the WISC-R verbal scale will correlate most highly with the BITCH, the entire (verbal plus performance) scale will be administered Ss, so as to take advantage of any possible interaction effects, as well as to use any additional data that may be followed up in future studies.

Following the test sessions, and at a later date, the general results of the study will be discussed with interested parents, teachers, and Ss. In every case, however, the anonymity of Ss will be insured through a coded numbering system. Confidentiality of the data, where Ss identity might be revealed, will also be protected; only this writer will hold lists containing both Ss name and number. All publications will only release the generalized results; at no time, either now or in the future will this writer release names or other specific data which could compromise Ss identity. It is not anticipated that any test or experimental procedure will be likely to produce harmful effects on Ss. Chapter III will

---

[38] Williams, Op. cit., p. 11.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

contain considerably greater depth of description regarding the procedures, sample, collection of data, and other pertinent information.

Analysis of Data

In testing the various hypotheses, appropriate parametric statistical measures, where applicable, will be employed. To be more specific, for Hypotheses I and II, the Pearson product-moment correlational procedures will be forthcoming, following determination of means and standard deviations for each set of data. It should be noted that a sample size of 60 (with 58df) requires an r of .25 or better for significance at the .05 level of confidence. A sample size of sixty appears to be adequate for purposes of statistical inference.[39] For Hypotheses III, the differentiation of social class represents an artificial dichotomy which departs from strictly parametric data; the possibility of either a Biserial correlational technique or Kendall's Tau will be considered. In exploring interaction effects involving WISC-R subscales, and the BITCH, analysis of variance procedures will be used to determine significant differences between sample means. In all cases, an Alpha level of .05 will be considered significant for purposes of this study.

---

[39] C. D. Hardyck; Z. F. Petrinovich; Statistics for the Behavioral Sciences, (Philadelphia, Pa.: Saunders, 1976), p. 308.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Limitations of Study

a. It is not anticipated that there is a significant difference between males and females in the target sample; accordingly this study will not hypothesize that such differences do, in fact, exist. Efforts will be made, however, to observe any differences in the sample. It is noteworthy that in spite of reported sex differences on a tangentially-related study, these investigators did not satisfactorily account for the variance, suggesting only that, "perhaps males exhibit more learning problems in the typical public school classroom than do females."[40] Without further specificity or offering of testable hypotheses, the sex differences have yet to be clarified. Since other studies (performed on the WISC-R) reveal no statistically significant sex-differentiated data, it is unlikely that such a relationship actually exists.

b. This study will attempt to test various hypotheses (previously cited) with a typical secondary school population. Although important data may well be yielded by selecting atypical (e.g., retarded, handicapped, disturbed) populations for future study, the selection of a normal population is considered important to avoid statistical regression effects. Normality in this case is defined by teacher judgments, grades, and continued attendance in school.

---

[40]Long and Anthony, Op. cit., p. 312.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

c. It is felt that examining population samples for differences on geographical or regional variables, particularly rural-urban discrepancies, is likewise an important dimension, but outside the scope of this study.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER II

REVIEW OF RELATED LITERATURE

Introduction

A quite extensive literature exists which closely relates to the nature and testing of human abilities; it has accumulated increasingly over the past century and can be subdivided into several fairly distinct periods. Paralleling the growth of ability testing has been the development of great controversy, with vast social and ethical implications which persist, and indeed have intensified to the present time. This chapter will address itself to the early theoretical development, the rise and spread of individual intelligence testing, the controversial implications, and its application to this study.

I. Early Theoretical Development

The work of Charles Darwin was of the utmost importance to the ongoing development of all studies relating to living organisms.[1] As O'Neill states:

---

[1]Charles Darwin, 'Brain, Sympathy and Survival," from Samuel Beck and Herman Molish, Reflexes to Intelligence, (Glencoe, Ill.: Free Press, 1959), pp. 8-16.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> "...The question of the survival value of various actions to the group and/or to the species was raised by his work. Of psychological importance was his concept that if individual differences among the various sub-human species relate to the differences in survival rate, individual differences in human beings may relate to the survival rate of the human species. These differences should be identifiable and measurable."[2]

Francis Galton was heavily influenced by Darwin's work. He was convinced that the majority of human traits were the result of heredity; these included both human abilities, as well as personality. In attempting to measure such abilities, various mental tests and standardization methodologies were developed by him. In Galton's era, the notion was paramount that:

> "...all knowledge comes through the senses. Accordingly, the conclusion was made that the person with the most acute senses would be the most knowledgeable. Because of this orientation, he developed a variety of physiological tests."[3]

The term "mental test" was first used by J. McKeen Catell in 1890.[4] The tests consisted primarily of samples of such human abilities as reaction time, color naming, and memory span. Galton had alluded to the concept of mental testing, but had not specifically formulated the term in an article

---

[2]Hugh Daniel O'Neill, Jr., Partial Validation of the Slossen Intelligence Test, Unpublished Doctoral Dissertation, (Norman, Okla., 1969), p. 8.

[3]O'Neill, Ibid., p. 9

[4]J. McKeen Cattell, "Mental Tests and Measurements," Mind, V. 15, (1890), pp. 373-380.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

published in 1883.[5] Binet and Henri (1895) appear to have formulated the working definition, stating that:

> "...mental testing methodology consists in the selection of a number of tasks designed to give detailed information on individual differences."[6]

It should be noted that the emphasis is on task-related measures which are designed to shed light upon individual differences. Terman indicates that succeeding definitions have been developed along parallel lines, but have attempted to differentiate even more clearly between the mental test and the psychological experiment.[7] In somewhat later collaboration with Simon, Binet created the first global measure of general intellectual ability in 1905.[8] O'Neill indicates that:

> "...Binet's work was of primary importance in that he...proceeded to investigate overall intellectual ability...he also utilized previous experience to develop his evidence. With Binet's work, the subsequent development of practical measures of intelligence was rapid."[9]

---

[5]Francis Galton, Inquiries Into Human Faculty and Its Development, (London: MacMillan, 1883).

[6]Alfred Binet and V. Henri, "La Psychologic Individuelle," Annee Psychologique, (1895), V. 2, pp. 411-463. Quoted in Louis Terman, "The Mental Test as a Psychological Method," Psychological Review, V. 31, (1924), pp. 93-117.

[7]Louis Terman, Loc. cit.

[8]Jerome Sattler, Assessment of Children's Intelligence, (Philadelphia, Pa.: Saunders, 1974), p. 87.

[9]O'Neill, Loc cit., p. 10.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

II. The Controversial Intelligence Test: Historical, Social, and Ethical Implications

Although Alfred Binet was instrumental in developing the first individual intelligence test, it remained for the Americans to encourage the testing of human abilities and to advocate its usefulness throughout all aspects of society.[10] Seemingly in society today we possess an unquenchable thirst for quantification; almost every magazine or popular periodical contains tests for the reader to self-administer, so as to determine how we think, function, interact, and feel about a variety of factors almost as intricate as the hand computers to which we are so currently dedicated. We are:

> "...inveterate and addicted stop watchers, slide rules, tabulators, counters, computers, and, above all, testers. We test aptitude, we test ability, we test achievement. Foremost, we test intelligence. We not only test it, we can reduce our assessment of it to precise numbers; our Intelligence Quotient, or the I.Q. So we like to think. So we have always thought..."[11]

To understand the spread of testing and its beguiling concepts of reducing complex human attributes, traits, and thoughts to quantities, numbers and rankings, we can turn to Hofstadter's penetrating analysis of the newly emergent American society with its emphasis on practicality, pragmatism, and organization.[12] Americans were out to "get the job

---

[10] Block and Dworkin, Op. cit.

[11] Jack Fincher, Human Intelligence, (New York, N.Y.: Putnam and Sons, 1976), p. 16.

[12] Richard Hofstadter, Social Darwinism in American Thought, Rev. Ed., (New York, N.Y.: Braziller, 1969).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

done." Focus on the esoteric or philosophical gave way to reduction of complexities to the simplest terms possible to accomplish the task. The more we knew about ourselves and the environment, the easier it was to analyze the fundamentals necessary to job completion. In those heady days when Americans infused with the doctrine of "manifest destiny" had established the United States "from sea to shining sea," the concept of "bigger and better" was given free rein. No job was too big; there seemed little that Americans couldn't do. After all, Americans had built the Panama Canal, developed the airplane, invented the electric light, and refined the egalitarian ideal to include the concept of extending freedom to all those whose former countries were happy to see leave. All an individual had to do was to understand himself, apply himself, and help himself--this being the case, wealth, happiness, and success were assured.

Today the intelligence test is over seventy years old, a "senior citizen in the growing American family of psychometrics."[13] It has shaped and, in turn, has been shaped by prevailing American social philosophy. It has exerted a powerful influence on the lives of the public, particularly school children whose placement and progress has been largely determined and continuously monitored by this instrument. The product of the intelligence test, the I.Q., has entered every stratum of our society and has, in many cases profoundly affected a person's view of himself, perhaps for life.

---

[13]Fincher, Loc. cit.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

What is this instrument and why should it hold such power over our lives and self-conceptions?

In reviewing the introduction and spread of ability testing, it can clearly be seen that a major tool to use in assisting society to improve itself, and indeed to help the individual to "better his existence," was felt to be the intelligence test. Cronbach has ably summarized much of the initial fervor and optimism with which this instrument was viewed and why this was the case.[14] He quotes Callahan who states:

> "...a principal tenet of the Progressive movement at the turn of the century was the power of social science to redirect and reshape society. Darwin and Comte, between them, had made the point that man and his institutions are subject to scientific, iconoclastic analysis...The American reformers expected factual analysis to free society from the ills that ranged from political corruption to prostitution, from despoilation of the environment to child labor...Efficiency and scientific management (as seen) in industry, were to be brought to social institutions."[15]

Mental testing had gained great notoriety following Galton's studies, as well as the introduction of the concept of mental testing of individual differences. The popularization of testing paralleled the development of the first large-scale group intelligence tests, growing out of the need for classification of Army recruits, with the advent of World War I.

---

[14]Lee J. Cronbach, "Five Decades of Public Controversy Over Mental Testing," American Psychologist, (January 1975), pp. 1-14.

[15]Cronbach, Ibid., p. 8.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> "...The mental test, fresh from its triumphs in the Army, promised to sort out pupils who should move fast, those who should move slow, those who should go to college, and those who should not... The tests would allow grouping by ability and would pick out the talented children for special encouragement...The tester's sorting process was to shield the child destined to be a worker from the rigors of an academic curriculum. Such a plan would reduce distaste for schooling, prevent failure, and retain him in school longer. Testers said the I. Q. was constant; hence to make decisions early was merciful and just."[16]

To place this argument in perspective, it is important to focus upon the Zeitgeist and the mood of the times. The movement for progressive reform under Theodore Roosevelt, combined with American pragmatism, set the stage for widespread public acceptance of the intelligence test. Even as the public mood favored testing, the seemingly perennial controversy over the extent to which mental abilities were inherited or were largely environmentally-determined, gained prominence. A review of the pertinent literature of those times, particularly the argument between Terman on one hand, and Lippman on the other, gives ample evidence of the vitriol which such debates can generate.[17] Terman's position regarding the innate qualities of intelligence versus Lippman's view that environment was all-important, foreshadowed a continuing controversy which, although waxing and waning in the public eye, was only waiting offstage to reappear when

---

[16]Cronbach, Ibid.

[17]Block and Dworkin, Op. cit., pp. 4-44.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the times favored a regeneration of the issue. In this instance, it must be stated that "out of sight" was definitely not "out of mind." Periodically, articles written on various aspects of this subject might be selected by editors for publication when such debate or dissent was "timely" or newsworthy. Thus in the early 1920's there was considerable sentiment against the continued and unrestricted flow of immigrants to these shores. Racism and ethnocentrism aimed at immigrants from Southern and Eastern Europe was active. One psychologist (Brigham), taking the Army Alpha test results for ethnic groups, published a book and several articles "proving" the case for innate racial and ethnic differences between these immigrant groupings and other segments of the European populace which had emigrated to the United States.[18] In support of his position, he indicated that the tabulation of results clearly showed that Army recruits with Italian or Polish backgrounds, for example, fared much more poorly in ability testing than men from Western European backgrounds (although variables involving education, social class status, and acculturation indices were not held constant).[19] With the passage of the much more restrictive immigration laws of 1924, which established quotas based on nationality and land of origin, the furor over ethnic differences (and supportive

---

[18] C. C. Brigham, "Value of Tests in Examinations of Immigrants," Industrial Psychology, V. I., (1926), pp. 412-417.

[19] C. C. Brigham, A Study of American Intelligence, (Princeton, N.J.: Princeton University Press, 1923).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

articles) quickly passed from the scene.

Although the heredity-environment issue was not necessarily in the public eye, it was the subject of continued controversy by psychologists and other professionals before and after World War II. Following the school desegration decisions of 1954 by the United States Supreme Court, a resurgence of public concern over racial differences (especially between Black and Caucasian groupings) on the subject of intelligence began to emerge. At issue were the social consequences inherent in classifying abilities along racial lines and from this data, predicting educational possibilities, job opportunities, and even societal rankings.[20] It was of vital concern, especially to Black minority groups involved with the overriding issue of securing equal rights and opportunities, that non-White populations not be labelled "intellectually inferior" on the basis of intelligence testing. Prior concern had already been generated by the publication of an extensive summary volume by Shuey. She studied the results of over fifty years of research between Negro and Caucasian groups on ability measures. Her conclusions were that supportive data obtained from over 380 studies involving 14,800 Negro Ss, "...all taken together, inevitably point to the presence of native differences between Negroes and Whites,

---

[20]Loelin, Lindzey and Spuhler, Op. cit., pp. 8 ff.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

as determined by intelligence tests."[21] In this respect, she is echoed by Garrett, who in the preface of her book says,

> "Dr. Shuey finds that at each level and under a variety of conditions, Negroes regularly score below Whites...We are forced to conclude that the regularity and consistency of these results strongly suggest a genetic basis for the differences. I believe that the weight of evidence (biological, historical, and social) supports this judgment."[22]

Even this definitive a statement did not, at the time, bring the opposition "big guns" to bear, although some minor flurries did result. Loelin indicates that hers was considered by the opposition to be largely a summary book, thus offering only indirect evidence and, "in any case, reviews with very different conclusions were available" (Klineburg, 1944; Dreger and Miller, 1968). Pettigrew has strongly attacked Shuey's and Garrett's conclusions citing the fact that the studies were poorly controlled, regionally biased (Southern United States), and replete with examiner bias and/or misconceptions.[23] After carefully dissecting and refuting their arguments, Pettigrew offers a 1961 statement by the Society for the Psychological Study of Social Issues, a division of the American Psychological Association, to wit:

> "...there are differences in intelligence test scores when one compares a random sample of Whites

---

[21]Audrey Shuey, The Testing of Negro Intelligence, (New York, N.Y.: Social Science Press, 1966), p. 521.

[22]Shuey, Ibid., p. viii.

[23]Thomas Pettigrew, A Profile of the Negro American, (Princeton, N.J.: Van Nostrand, 1964), pp. 100-135.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> and Negroes. What is equally clear is that no evidence exists that leads to the conclusion that such differences are innate. Quite to the contrary, the evidence points overwhelmingly to the fact that when one compares Negroes and Whites of comparable cultural and educational background, differences in intelligence diminish markedly; the more comparable the background, the less the difference. There is no direct evidence that supports the view that there is an innate difference between members of different racial groups...We regret that Professor Garrett feels that his colleagues are foisting an 'equalitarian dogma' on the public. There is no question of dogma involved. Evidence speaks for itself and it casts serious doubt on the conclusion that there is any innate inequality in intelligence in different racial groups."[24]

Following the publication of Shuey's book, the stage was set for a major explosion.

Into this arena, a major article by Arthur Jensen was published by the Harvard Educational Review.[25] Jensen, a well-known Educational Psychologist, stated that the vast weight of evidence on Negro/White intellectual differences strongly supported the case for genetic factors determining an individual's ability. Not only does inheritance account for about eighty percent of the individual's intelligence, according to Jensen, but the continued discrepancy of about fifteen I.Q. points favoring Caucasians over Blacks implied, at the very least, an innate inferiority in ability for the latter group.[26]

---

[24] Pettigrew, Ibid., pp. 133-134.

[25] Arthur Jensen, "How Much Can We Boost I.Q. and Scholastic Achievement?" Harvard Educational Review, V. 39, (1969), pp. 1-123.

[26] Arthur Jensen, Genetics and Education, (New York, N.Y.: Harper and Row, 1972), pp. 157-172.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In fairness to Jensen, several points should be made; first, he was genuinely concerned over how to compensate the Black child through appropriate educational programs; second, he indicated that although the Black genotype was intellectually lower, the phenotype can vary by as much as twenty-five I.Q. points if raised in an enriching and stimulating environment.[27] Jensen cited such works as Shuey, as well as Burt's celebrated studies of identical twins reared apart,[28] in addition to his own research, to support his position.

In order to more clearly understand the furor this article generated, one must again focus on the social change which had been gathering momentum, especially for minority groups, in the decade between 1960 and 1970. Not only were individual differences and "talent" rewarded, but the value of high test scores became critical. If one fared badly on college entrance tests, no draft exemption (from the Vietnam conflict, for example) was available. Test scores became major discriminators in preventing an individual from participating in the "good life" in terms of college and career opportunities. The more test results became the determinants of one's future, the more they aroused anxiety and public concern. This was particularly true among minorities where greater social opportunities than ever before were just

---

[27]Jensen, Ibid., pp. 179-193.

[28]Cyril Burt, "The Genetic Determination of Differences in Intelligence: A Study of Monozygotic Twins Reared Together and Apart," British Journal of Psychology, V. 57, (1966), pp. 137-153.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

starting to appear in the United States. Hard-won gains in Black America, the result of countless "sit-ins," marches, rallies and (occasionally) riots, had given Blacks, perhaps for the first time, a taste of what the future could be for them; they were not going to let this moment slip away.

Along with the public (if sometimes grudging) acceptance of minority gains, however, a "worm appeared in the apple." The Head Start Program, which was designed to help equalize environmental differences between pre-school age children of different social classes, racial and ethnic backgrounds, did not demonstrate the gains that its supporters had so widely touted.[29] If such a program (however hastily produced), displayed little long-term benefit, especially with regard to Black/White groupings, might there be the possibility that Blacks were, indeed, intellectually inferior by virtue of genetic endowment, to the Caucasian population? This fear surfaced in the late 1960's at precisely the time that Jensen's article was published. If a well-known psychologist from a prestigious university writes such an article, and if it is published in as prestigious a journal as the Harvard Educational Review, minority suspicions (always ready to react to the spectre of bias) had found the perfect catalyst. As Cronbach indicates,

> "...For perspective, the Jensen affair must be seen along side other controversies. Particularly important to the story is the role of the media in shaping controversies...Nor are the

---

[29]Jensen, Op. cit., pp. 60 ff.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> media to be seen as independent agents; what catches fire depends, at any time, on the way public opinion is blowing."[30]

Immediately upon publication of Jensen's article in the Harvard Educational Review, a furor erupted which created wave upon wave of dissension, not only among academic ranks, but also in the government. One of the advisors to President Nixon was Daniel Moynihan who was accused of not only distributing Dr. Jensen's article to the Nixon cabinet as "must" reading, but also of biased advocacy of this "racist" (i.e., Jensenian) point of view.[31] Moynihan, now the Junior Senator from New York, felt compelled to write a rejoinder denying this allegation. He stated that although he was strongly committed to an environmentalist position regarding the distribution of intelligence by race, he knew what Jensen was going through, stating, "I got the same treatment for almost exactly the opposite hypothesis!"

The list of prominent figures from the academic world who either agreed or disagreed with Jensen reads like a "who's who" in Psychology, Education, Genetics, and associated fields. Psychologists (Anastasi, 1971; Hunt, 1971; Burt, 1972: S. S. Stevens, 1971; Eysenck, 1971), Geneticists (Dobzhansky, 1970),

---

[30]Cronbach, Op. cit., p. 11.

[31]E. Alfert, "Comment On: The Promotion of Prejudice," Journal of Social Issues, V. 25(4), (1969), pp. 206-211.

[32]Daniel Moynihan, "Comment: Jensen Not 'Must Reading' in the Nixon Cabinet," Journal of Social Issues, V. 26(2), (1970), pp. 191-192.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Sociologists (Bennett, 1971), Physicists (Shockley, 1972), are but a very few of the contributors to the voluminous literature sparked by Jensen's 1969 publication in the Harvard Educational Review. If the academicians rallied to the call, the media transmitted a message to the public which touched off major public debate which has continued to this very day. Subsequently, newspapers and magazines, including The New York Times,[33] San Francisco Chronicle,[34] and the Saturday Review,[35] again, to name but a few, helped to polarize public opinion. In short, the I.Q. controversy has been one of the liveliest and most "saleable" public controversies from 1969 to date. And what was the issue publicized by the press? Basically, it dealt with a "popularization" of the notion that Blacks are genetically inferior to Caucasians regarding intelligence and, in view of this factor, consigned to menial occupations and social positions. A "permanent lower class of the unintelligent...which runs in the genes like rotten teeth."[36]

Misquotes and entire phrases were lifted out-of-context in condemning Jensen and advocates of his position,

---

[33]Lee Edson, "Jensinism, n., The Theory That I.Q. Is Largely Determined by the Genes," New York Times Magazine, (8-31-69).

[34]Jim Wood, "The I.Q. Bias Battle, San Francisco Sun-Day Chronicle, (9-24-72).

[35]Robert McQueen, "Larry: Case History of a Mistake," Saturday Review, (9-12-70).

[36]Harvard University Action Group, Quoted in Richard Herrnstein, I.Q. In the Meritocracy, (Boston, Mass.: Atlantic-Little-Brown, 1973), p. 22.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

especially Herrnstein.[37] More importantly, in terms of objective research, adherents of genetic intellectual primacy were automatically branded "racist" and were condemned by a vast portion of their colleagues (as well as the public) who sought to claim their purity of research while pointing the finger of bias at Jensen, et al. Genetic advocates were denied grants, funding and forums;[38] they were personally threatened with bodily harm by activist groups.[39]

Running throughout the entire controversy was, of course, the fate of the "harbringer of bad tidings," the intelligence test. This instrument was at fault--it "reduced" human potential to an I.Q. score and was viciously attacked as being an inadequate measure of ability, especially as applied to a Black population. Citing inappropriate or misleading normative studies, a number of standardized tests, particularly the Wechsler and Stanford-Binet Scales, were deemed unfair and biased against minorities. The Association of Black Psychologists (ABP) further charged that such instruments were used by school districts to "label" and wrongly place Black children in classes for the mentally retarded, when in fact no such label was appropriate or germane.

In 1972, the State of California and the San Francisco Unified School District were sued by the San Francisco Bay

---

[37]Herrnstein, Ibid., pp. 3-59.

[38]Jensen, Op. cit., pp. 1-67.

[39]Herrnstein, Loc. cit.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Area Association of Black Psychologists on behalf of "Larry P." and six Black school children which this group claims were inappropriately placed in Educable Mentally Retarded special education classes in the district.[40] A moratorium on such testing was demanded by the ABP and, in 1972, an injunction against the using of the individual intelligence test in determining minority retardation was granted by Judge Robert Peckham in whose court the "Larry P." case is currently being heard. Moreover, a moratorium on the use of group intelligence tests has been in effect in California schools since 1972. In the interim period since the filing of the initial injunction against intelligence testing and classification of minority children, a number of articles citing the ABP viewpoint have appeared in popular print.[41] "Semiprofessional" publications equating minority intelligence testing with "Black intellectual genocide" and the "silent mugging" of the Black community have also abounded.[42]

Today we have come full circle to the very dilemma which initiated large-scale intelligence testing, and for which such testing was purported to be the answer. Rather than "opening doors" and espousing the egalitarian ideal, however, intelligence tests are frequently seen by the lay

---

[40] Larry P., et al, vs. Wilson Riles, et al., #C 71-2270 RFP, U. S. District Court: Northern District of California, (June 1972).

[41] Wood, Op. cit.

[42] Robert Williams, "The Silent Mugging of the Black Community," Psychology Today, (May 1974).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

public as "discriminative devices" which are biased against minority groups. Ebel has recently indicated that:

> "...Many of the popular articles critical of educational testing that have appeared in recent years do not reflect a very adequate understanding of educational testing, or a very thoughtful, unbiased consideration of its social consequences...What appears in print often seems to be only an elaboration and documentation of prejudices and preconceptions, supported by atypical anecdotes and purposefully selected quotations. Educational testing has not fared very well in these articles."[43]

Although testing opponents have given the intelligence test a very "bad press," its defenders state that tests cannot be blamed for indicating the deficiencies of the child. Pointing the finger at ability testing and advocating its abolition is not unlike the practices of ancient kings who used to "kill the messenger that brought the bad news."[44] Dyer states:

> "...Some of our brasher critics have argued that, since tests are so widely misused, they do constitute a menace to sound education and should be abolished. This argument is specious. It is the same as saying that automobiles should be abolished because they are a menace to human life when reckless drivers are at the wheel. Or it is the same as saying that teachers should be abolished because too many of them make psychometric hash out of marks and test scores."[45]

---

[43] R. L. Ebel, "The Social Consequences of Educational Testing," Proceedings of the 1963 Invitational Conference on Testing Problems, (Princeton, N.J.: Educational Testing Service, 1963), pp. 130-143.

[44] H. S. Dyer, "Is Testing a Menace to Education?" N.Y. State Education XLIX (October 1961), pp. 16-19. In, Readings in Educational and Psychological Measurement, ed., C. I. Chase and H. G. Ludlow, (Boston, Mass.: Houghton-Mifflin, 1966), pp. 40-45.

[45] Dyer, Ibid., p. 45.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Indeed, those who would abolish ability testing must understand that its elimination would not cause the disappearance of the needs that testing presently serves. Other means to appraise performance would have to be provided. Once objective and standardized comparisons are abandoned, it is very likely that more subjective alternatives would emerge, thus enhancing the possibility of even greater bias and discrimination.[46] Ebel has also concluded that,

> "...the social consequences of not testing...are potentially far more harmful than any possible adverse consequences of testing. Important curriculum decisions and methodology...would be made less on the basis of solid evidence and more on the basis of prejudice or caprice...distinctions between competence and incompetence would become more difficult to discern. Educational opportunities would be extended less on the basis of aptitude and merit and more on the basis of ancestry and influence; social class barriers would become less permeable."[47]

A test is a series of controlled observations, and as Samuda indicates,

> "...the burden or responsibility shifts to the test user, whose duty it is to be alert to and elminiate unfair circumstances, biased items, or inappropriate questions. To say a test is fair or unfair is inaccurate, as it is the particular use of the instrument which may be fair or unfair."[48]

In summarizing the current testing controversy, Samuda feels that although attacks on the testing industry will undoubtedly continue, this will probably not result in the abandonment of

---

[46] S. Messick and S. Anderson, "Educational Testing: Individual Development and Social Responsibility," Counseling Psychologist, V. 2, (1970), pp. 80-88.

[47] Ebel, Op. cit., p. 143.

[48] Samuda, Op. cit., p. 13.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

objective or standardized testing. Nor will attacks on individual school districts eliminate the use of such instruments. The intelligence test, for all its problems, is a useful and well-entrenched device in American society. Changes are, however, mandatory if they (tests) are to function in useful societal capacities.

Fishman and his associates have endorsed a similar concept in drafting the "Guidelines for Testing Minority Group Children."[49] They advocate not the elimination of those instruments (ability tests) which are, "among the most important evaluative and prognostic tools that educators have at their disposal."[50] They focused upon three major problems when dealing with minority group testing:

1. Tests may lack reliable differentiation in the range of minority group scores, which tend to cluster at the lower tails of the curve.

2. Tests may not predict for the minority child what they would for the middle-class Caucasian child.

3. Test results should be carefully interpreted by professionally trained personnel who are quite familiar with the socio-cultural background of the group being tested.

The "Guidelines" concluded with the statement"

> "Many comparisons depend upon tests, but they also depend upon our intelligence, our good will, and our responsibility to make the proper comparison at the proper time and to undertake proper remedial and compensatory actions as a result. The misuse of tests

---

[49] J. Fishman, et al., "Guidelines for Testing Minority Group Children," Journal of Social Issues, Supplement, V. 20, (1964), pp. 129-145.

[50] J. Fishman, et al., Ibid., p. 143.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> with minority group children, or in any situation, is a serious breach of professional ethics. Their proper use is a sign of professional and personal maturity."[51]

Regardless of the ultimate outcome of the testing controversy, an issue that looms as large in its effect upon the future of research deals with the ethical problems raised by the furious reaction within the academic and research community to Jensen's and Herrnstein's articles. Jensen, in discusing the ethical implications of his research, states:

> "...but the most frequently heard objection to further research into human genetics, particularly research into the genetics of behavioral characteristics, is that the knowledge gained might be misused. I agree. Knowledge also, however, makes possible greater freedom of choice. It is a necessary condition for human freedom in the fullest sense. I therefore completely reject the idea that we should cease to discover, to invent and to know (in the scientific meaning of the term) merely because what we find could be misunderstood, misused, or put to evil and inhumane ends."[52]

In a free society, he continues, it is imperative that the scientist publish results, taking pains to competently conduct research and report his observations in an unbiased and accurate fashion; to do otherwise is to serve political ends which can rebound and ultimately do more harm than good. Jensen cites examples of such state-serving research that abounded in more extreme forms in Nazi Germany or Stalin's Russia. Even milder instances of research suppression (as in the case of genetic factors and their influence on human

---

[51] J. Fishman, et al., Ibid., p. 144.

[52] Jensen, Op. cit., p. 327.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ability) result in potentially dangerous consequences to future generation. He quotes Ingle, a Geneticist, who states, "If there are important average differences in genetic potential for intelligence between Negroes and non-Negroes, it may be that one necessary means for Negroes to achieve true equality is biological."[53]

Jensen feels that the dysgenic trends in current society (with the racial and social class differential in birthrates), could easily serve to widen a gap between Blacks and Caucasians, with enormous future social consequences. He further believes that the failure of the United States government to explore these possible consequences through funding research in this area, "may not be unethical--but it is, I believe short-sighted, socially irresponsible, and inhumane."[54] Finally, Jensen greatly resents being cast as the "bad guy" merely because he is a researcher in the genetic component of intelligence in the population; he says:

> "...The simple-minded morality play in which I have been wittingly or unwittingly cast in the role of villain has presented the issue of ethics as if ethical behavior were the sole possession of the environmental dogmatists, and as if those of us who would suggest looking into genetic factors were ethical or moral pariahs!...In my view, society will benefit most if scientists treat these problems in the spirit of scientific inquiry rather than as a battlefield upon which one or another

---

[53] D. J. Ingle, "Editorial: The Need to Study Biological Differences Among Racial Groups: Moral Issues," Perspectives in Biology and Medicine, V. 10, (1967), pp. 497-499.

[54] Jensen, Op. cit., p. 131.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

preordained ideology may seemingly triumph."[55]

A similar theme is discussed in a thought-provoking article by Scriven.[56] He raises a number of issues pertaining to American education and the highly negative and often vicious reaction of the academic world to Jensen's research. Some of the most telling arguments against those who attack Jensen lie in the fact that many of them argue illogically or do not possess elemental but necessary facts to muster in refuting his case. This, Scriven feels, is an indictment against the university training and competence of many of Jensen's detractors. In separating the truly arguable issues from the emotional (and often illogical) ones, Scriven refutes the position of some that it is incorrect to publish results that can be mininterpreted by self-serving or misinformed individuals. He makes the point that research on racial differences is indeed valid and not to be confused with or labelled racism; the two are not synonomous. Jensen, by no stretch of the imagination, judging from his published works is racist, except, "in the sense that his work can be (mis-) used by racists to support their case. If that kind of redefinition is allowed, then the revolutionaries are reactionaries, since their activities provide support for reac-

---

[55]Jensen, Ibid., p. 332.

[56]Michael Scriven, "The Values of the Academy (Moral Issues for American Educational Research Arising from the Jensen Case)," Review of Educational Research, V. 40, (1970), pp.541-549.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

tionary programs."[57] Scriven feels that Jensen has asked important questions which, if intelligently pursued, could lead to important restructuring of academic goals, which is after all, an overriding value for our society. He says:

> "... In general, it is worth remembering that the standard of criticism here is not whether fifty critics can pick some holes in a hundred page survey article covering several hundred references--each critic knows this would happen to him if he was the author--but whether the criticism destroys the contribution of the article. The attendant furor in the Jensen affair has focused all attention on one point in it; but even on that one point, and after all the criticism, Jensen's case is, I believe, not only well within the boundaries of legitimate professional interpretation, but an extremely important position to consider."[58]

The Problem of Minority Group Testing and Its Application to This Study.

The standardized or "norm-referenced" intelligence test has long been thought to provide less predictive value for those children who are in any way distanced or alienated from the cultural mainstream. Such children might be verbally disadvantaged, handicapped, or simply not exposed to the informational or experiential base upon which norm-referenced tests draw. Cattell, for example, wished to introduce an instrument which would control for the imposition of the culture upon the individual.[59] Davis and Eells also felt the need

---

[57] Scriven, Ibid., p. 547.

[58] Scriven, Ibid., p. 549.

[59] Raymond B. Cattell, "A Culture-Free Intelligence Test," Journal of Educational Psychology, V. 31, (1940), pp. 161-179.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

to produce an instrument which would be free of cultural influences. They theorized that when such factors are eliminated, no truly important differences would exist between the basic intelligence of children from diverse social class or ethnic backgrounds.[60] The culture-free approach grew out of the notion that, "native intelligence lies buried in pure form deep in the individual and needs only to be uncovered by ingenious mining methods."[61] It became increasingly clear, however, that the attempts to dichotomize heredity and the environment were either fruitless, or led to the confounding of those variables. Jules Henry, in presenting Rorschach plates to Pilaga (Amazonian jungle) Indians found that so-called "universal" responses to the inkblots were intimately, and heavily influenced by the social matrix in which the child matured.[62] As it was the case with visual perception, so also was it true with linguistic process as Sapir[63] and Whorf[64] discovered. At a somewhat later date, Hallowell, in research-

---

[60] A. Davis; K. Eells, Davis-Eells Test of General Intelligence or Problem-Solving Ability Manual, (New York, N.Y.: World Book Co., 1953).

[61] A. G. Wesman, "Intelligence Testing," American Psychologist, V. 27, (1968), pp. 267-274.

[62] Jules Henry, "Rorschach Technique in Primiitve Cultures," American Journal of Orthopsychiatry, V. 11, (1941), pp. 230-234.

[63] E. Sapir, "Conceptual Categories in Primitive Languages," Science, V. 74, (1931), p. 578.

[64] B. L. Whorf, "Science and Linguistics," In E. L. Hartley and T. M. Newcomb, (eds.), Readings in Social Psychology, (New York, N.Y.: Holt, 1947).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ing the response pattern of Canadian Indians, essentially validated Henry's observations.[65] Anastasi has stated,

> "...heredity and environmental factors interact at all stages in the organism's development, from conception to death, and...their effects are inextricably intertwined in the resultant behavior."[66]

Thus the emphasis shifted from the "culture-free" concept toward the development of "culture-fair" tests. Since cultural effects couldn't be eliminated, test creators attempted to, "keep cultural differences from permeating the tests by selecting only those experiences, knowledge, and skills common to different cultures."[67] Basically, these instruments tend to minimize any factors which are believed to penalize the response of lower-class children (e.g., item content, verbal content, speed). They are basically nonverbal, with materials presented with either a minimum of language, or with gestures, where appropriate. The items are likewise selected on the basis of their "universal" quality and consist of pictures, puzzles, drawings, or diagrams. In the item selection process, only those stimuli which "motivate equally all the groups to be tested"[68] are retained.

---

[65] A. I. Hallowell, "'Popular' Responses and Cultural Differences: An Analysis Based on Frequencies in a Group of American Indian Subjects," Rorschach Research Exchange, V. 9, (1945), pp. 153-168.

[66] Anne Anastasi, as quoted in Gronlund, N. E., Measurement and Evaluation in Teaching, 2nd Ed., (New York, N.Y.: MacMillan, 1971), p. 280.

[67] Samuda, Op. cit., p. 134.

[68] Eells, et al., Intelligence and Cultural Differences, (Chicago, University of Chicago Press, 1951), p. 23.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Some of the more prominent examples of culture-free and culture fair ability tests are the:

1. Cattell Culture-Free Intelligence Test
2. Davis-Eells Games Test
3. Raven Progressive Matricies
4. Leiter International Performance Scale
5. Goodenough-Harris Draw-A-Man Test

The available literature and research studies on the above instruments generally show minimal predictive or concurrent validation, that would support thier Raison d'Etre. As Samuda indicates:

> "it is the consensual opinion of psychometricians and psychologists that culture-free or culture-fair tests have proved disappointing and have fallen short of their goals, for minority students have been shown to perform, if not more poorly, at least just as badly as they do on conventional intelligence measures."[69]

The Culture-Specific Testing Movement

The underlying philosophy of culture-free or culture-fair ability tests views the disadvantaged and/or minority child from a cultural deprivation approach. Thus minority intelligence is explained as a deviation from the prevailing cultural norm which must be controlled or "equalized" throughout the test materials. Samuda states that the minority child, "having assumed successively the characteristics of the genetically deficient model and the culturally disadvantaged or deprived model...is now being described in the light of the culturally different model."[70]

---

[69]Samuda, Op. cit., p. 142.

[70]Samuda, Ibid., p. 17.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

One such approach which recognizes the child in terms of his cultural specificity, has emerged from the work of Jane Mercer.[71] She compares every child, as viewed from within his or her individualized cultural background. Group mem-membership is determined on the basis of familial and cultural characteristics which significantly differentiate between the particular subculture and the prevailing cultural norms of the larger community. She has applied her measurements to Mexican-American and Black groups, with results found to be at least initially encouraging. Essentially, this procedure consists of a battery of tests designed to supplement with their respective "strengths" the "weaknesses" that each may have in certain areas. She terms this approach a multidimensional one which includes measures of physical develop-ment, "adaptive behavior," and individualized student assessment. By utilizing norms for the child's physical, social and intellectual development, as compared with tables for Black, Hispanic and Caucasian children between five and eleven years of age, the learning potential of each child can be determined.[72] Her research indicates that culturally-specific adaptive behavior scales yield evidence in noncognitive areas which, taken together with appropriately-normed intelligence tests,

---

[71]Jane Mercer and June Lewis, System of Multi-Cultural Pluralistic Assessment (SOMPA), Technical Manual, Psychological Corporation (1977).

[72]Jane Mercer and June Lewis, SOMPA Test Kit Cover Letter, New York, N.Y.: Psychological Corporation, (1977).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

"provide a unique means of estimating learning potential that may be masked by sociocultural and health factors."[73] It should be noted that in addition to standard weight, growth, dexterity, and perceptual measures, Mercer employs the revised Wechsler Intelligence Scale for Children as her intellectual assessment device. In her System of Multi-cultural Pluralistic Assessment (SOMPA), she indicates:

> "...In connection with these measures (Adaptive Behavior and Intelligence) the child's welfare can be more seriously harmed by an underestimate of his or her strengths than by an overestimate of them... Interpretation of the Estimated Learning Potential (ELP) needs to be consistent with the assumptions of the pluralistic model. Thus it is a more serious error to underestimate the child's potential. In most cases, children having combined ELP scores of 85 or higher will be able to profit from remedial or other special instruction."[74]

Although, as previously indicated, the initial results look promising, Mercer's SOMPA is barely out of the pilot and research stage; it remains to be proven in academic and clinical usage.

Finally, the culture-specific movement should be viewed as a response to what their creators have felt was a built-in "bias" in norm-referenced or standardized intelligence tests. Apart from Mercer's approach, two culture-specific ability tests have gained some notoriety, while one of these two has succeeded in achieving some prominence. Adrian Dove and Robert Williams have each designed instruments which are intentionally biased toward the cultural experience of Black popu-

---

[73]Mercer and Lewis, Ibid., p. 2.

[74]Mercer and Lewis, Ibid., p. 2.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

lations. The DOVE Counterbalance General Intelligence Test was published in 1966 for tenth grade to adult populations. It is composed of questions which the author states would be answered reasonably quickly by disadvantaged Black adolescents and adults, although White middle-class groups would face considerable difficulty in correctly responding to the majority of items.[75] This instrument is sparsely documented and reported in the literature. Samuda and Williams mention the DOVE in passing, or in referring to his own test, in the latter case. The most recent publication (Oakland, 1977) does not discuss either the DOVE or give any prior data regarding research studies that have been initiated which use the test.

In contrast, Robert Williams' Black Intelligence Test of Cultural Homogeneity (BITCH) originally was created as a reaction to what its creator felt was a dehumanization of Black children through the medium of "biased" testing. Williams stated that majority children would do quite poorly on a device sensitive only to Black populations, and, within this article, several test items were published relative to what was then entitled the "Black Intelligence Tests Counterbalanced for Honkies."[76] From this beginning, spurred primarily by William's reaction to the genetic deficit theory of

---

[75]Adrian Dove, Dove Counterbalance General Intelligence Test, (1966), Adrian Dove, as mentioned in Compendium of Tests, Ronald Samuda, Op. cit., p. 182.

[76]Robert Williams, "Danger: Testing and Dehumanizing Black Children," Clinical Child Psychology Newsletter, V. IX, #2, (1970).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Black intelligence, has come the refinement, development, and standardization of what is probably the most prominent instrument of its type in current test literature.[77] Although others have partially developed culture-specific instruments for use with Black[78] or Mexican American populations,[79] the actual tests are either not available for review, or do not possess the local standardization base from which generalizations or comparisons can be drawn. The specific nature of the BITCH test will be described more fully in Chapter III of this paper.

Criterion-Referenced Testing

In completing this literature review, mention must be made of still another alternative to norm-referenced ability testing. Most prior test practices have been grounded in the development of national standards so that localized comparisons can be drawn. A different method focuses upon interpreting achievement by,

> "...describing in behavioral (or performance) terms the student's performances regarding a particular instructional objective without reference to the level of performance of other members of the group. The level of performance accepted as satisfactory is usually predetermined or even stated as part of each instructional objective. Thus the specific criterion behavior provides an absolute standard

---

[77] Oakland, Op. cit., p. 15.

[78] James Boone and Vincent Adesso, "Racial Differences on a Black Intelligence Test," Journal of Negro Education, (Fall 1974), pp. 429-435.

[79] C. C. Ortiz and G. Ball, "The Enchilada Test," Institute for Personal Effectiveness in Children, (1972).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> against which to compare an individual's achievement. Since a criterion standard rather than relative position in a norm group is used for describing test performance, such interpretations are called criterion-referenced. The design and construction of criterion-referenced tests, then, are directed toward obtaining measures of achievement that can be expressed directly in terms of student performance on clearly specified educational tasks."[80]

The basic concept of criterion-referenced testing antedates the study of psychometrics by millennia, since it was the teaching method used by the Greeks and Romans, as well as throughout the middle ages.[81] According to Samuda, use of percentage-mastery grade determination and reportage systems by European universities represent this type of approach.[82] Performance work measures used in industry are also of this genre.[83] Both Oakland[84] and Samuda[85] are lavish in their praise of this approach. However, since criterion-referenced tests depend upon an absolute, as opposed to a relative (norm-based) standard, the development of a good criterion requires strict delineation of objectives along lines discussed by Boehm.[86]

---

[80]Oakland, Loc. cit., p. 16.

[81]Samuda, Op. cit., p. 147.

[82]Samuda, Ibid.

[83]R. Jackson, "Developing Criterion-Referenced Tests," ERIC Clearinghouse on Tests, Measurement and Evaluation, (Princeton, N.J.: Educational Testing Service, 1971).

[84]Oakland, Op. cit., p. 16.

[85]Samuda, Op. cit., p. 147 ff.

[86]Ann Boehm, "Criteria-Referenced Assessment for the Teacher," Teacher College Record, V. 75, (1973), pp. 117-126.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

She enumerated the following:

1. Who determines the objectives?
2. Who sets the behavioral criterion levels?
3. Do test items accurately reflect the behavioral criteria?
4. What constitutes a sufficient sample of criterion levels?
5. Do the test scores obtained describe an individual's response pattern?[87]

Although the criterion-referenced approach may prove a potential boon to educators and scientists, it is still in its infancy with many problems yet remaining to be resolved; these include development of validity and reliability measures, appropriate and objective standards, and computer programming difficulties.[88]

Summary

In surveying this highly controversial field, emphasis has been given to the rise of the ability test, from its earliest foundations to its development and expansion throughout all aspects of our society. The use and misuse of the intelligence test has been cited, along with the publishing of its results, as applied to various segments and groups within our nation. The cyclical nature of social trends must be strongly considered whenever applications or conclusions based upon test results are enumerated, for their moral and ethical implications are great and tend to polarize scientists and academicians into opposing camps. Such adversarial relation-

---

[87] Boehm, Ibid., p. 120.

[88] Samuda, Loc. cit., p. 151.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ships are detrimental to objective scientific inquiry which is much to be preferred over biased or subjective approaches. Finally, the development of alternative methodologies to standardized intelligence tests, and their psychometric applications have been considered. The evolution and evaluation of various instruments and their comparison to traditional ability test methods have also been considered. The theory and application of one such instrument (Black Intelligence Test of Cultural Homogeneity), as compared to an appropriately-normed, standardized intelligence test, forms the major thrust of this study.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER III

DESIGN OF THE STUDY

Introduction

An interpretation of the findings of this study requires as comprehensive an understanding of the sample, the sources of data, and methodologies employed in this investigation as possible. This chapter, therefore, provides information regarding the selection of the sample, the description of the test instruments, and the methods by which the subjects were measured.

The Sample

The sample consisted entirely of Black adolescents residing in the San Francisco Bay Area. It was felt that by drawing a sample from an exclusively Black population, interaction effects occurring from interracial factors could be eliminated. Thus the results obtained from such a sample would be a purer measure of what is inherent within the test instruments themselves and any confounding factors produced by test-related racial differences would be controlled. Sixty-five (65) subjects, ranging in age from thirteen and one-half years (13 1/2) to sixteen years, eleven months (16-11) were used in the study. This range fulfilled the norming criteria for all tests employed. The WISC-R, for example, obtained

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

data from a standardization population age six (6) to sixteen years, eleven months, thirty days (16-11-30).[1] With regard to the age range for the BITCH, Williams notes that the test "is to be used primarily for adolescents and adults."[2] The California Achievement Test, level five, was designed to "provice a machine-scorable format for grades 1.5 through 12..."[3] Although this particular grade-range is extensive, the level five version of test concentrates on standardization norms from ninth through twelfth grades. This specific achievement range comprised the age and grade norms of the vast majority of the sample employed in this study.[4]

All of the Ss were obtained from secondary schools within the urban San Francisco Bay region, specifically the cities of Berkeley, Oakland and San Francisco, California between the dates of January to March of 1978. It was decided to use only those students who were described as average by teachers and counselors, in terms of academic grades and school attendance. It was felt that in this way a typical sample could be secured, so as to avoid any unwanted regression effects caused by the utilization of atypical Ss, which might tend to skew population parameters. Thus in obtaining the study

---

[1]Wechsler, Op. cit., p. 10.

[2]Williams, Cover Letter for BITCH Test Kit, p. 1.

[3]Ernest Tiegs and Willis Clark, California Achievement Test Manual, 1970 Ed., Level 5, Form A, Monterey, CA.: McGraw Hill, (1970), p. 5.

[4]Tiegs and Clark, Ibid., pp. 50-77.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

sample, teachers and counselors were requested to recruit or refer for consideration only those students that in their judgment and in terms of school attendance, were considered typical with respect to age/grade level/achievement.

Of the sixty-five Ss included in the study, thirty-one were male and thirty-four were female. Although no specific hypothesis was formulated regarding any sex differences, it had previously been reported in a comparison study of retarded Ss using the BITCH and WISC that such differences existed.[5] Accordingly, the subdivisions of the sample into an equal number of males and females could yield data on sex-related differences, if indeed such differentials existed within the sample studied.

The determination of social class differences coupled with the need to provide for clear-cut delineation between lower and middle social class status, necessitated a careful review of the literature regarding this variable. Most of the current sociological thought on this subject, as typified by several recent studies by Burnes[6] and Mach Erbe[7] use indices of income, occupation, and education of parent. Regarding this, Burnes states,

> "...SES was determined by the occupation of the head of household. Lower class status included

---

[5]Long and Anthony, Op. cit., p. 312.

[6]Kay Burnes, "Patterns of WISC Scores for Children of Two Socioeconomic Classes and Races, Child Development, V. 41, (1970), pp. 493-499.

[7]Brigette Mach Erbe, "Race and Socioeconomic Segregation," American Sociological Review, V. 40, (1975), pp. 801-812.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> unskilled work (laborer and domestic), unemployed status of the household head, and homes in which no father was present and the family depended solely on welfare support...Upper/middle class SES was defined by high-status professional occupations and important business positions. These included physician, owner of a large business, lawyer, etc. This group also included cases where the head of household was engaged in professional employment, e.g., social work, college teaching."[8]

Along this line, Mach Erbe indicates that, "...those factors determining SES among Blacks were determined by utilizing variables based on income, occupation, and education of parent."[9] Although as previously mentioned, Kahl points to the fallacy of giving primacy to variables of income in determining social class status[10] and Kohn eschews the use of income as a primary discriminator in assigning social status, [11] both investigators use income as but one of their stratification criteria. Also included in their analyses were variables relating to occupation and education and, in the case of Kahl's analysis, variables reflecting residential data (home type and location).[12] One of the more precise and oft-quoted stratification scales was prepared by Meier and Bell in their research involving SES differences and the achievement of life goals. [13] As this

---

[9]Mach Erbe, Op. cit., p. 803.

[10]J. Kahl and J. Davis, "A Comparison of Indexes of Socioeconomic Status, American Sociological Review, V. 20, (1955). pp. 317-325.

[11]M. L. Kohn, Class and Conformity, Ill.: Dorsey Press, (1970).

[12]Kahl, Op. cit., p. 319.

[13]Dorothy Meier and Wendell Bell, "Anomia and Differential Access to the Achievement of Life Goals," American Sociological Review, V. 24, (1959), pp. 189-202.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

index was a primary differentiator of social class, as used in this study, it consists of the following:

TABLE I

The Meier-Bell SES Scale

Income	Score
Under $1,000	1
$1,000-$1,999	2
$2,000-$2,999	3
$3,000-$3,999	4
$4,000-$5,999	5
$6,000-$9,999	6
$10,000-$16,999	7
$17,000-$29,999	8
$30,000 and over	9

Occupation	Score
Laborers	1
Service Workers	2
Operatives & Kindred	3
Farmers & Farm Managers	4
Craftsmen, Foremen & Kindred	5
Sales, Clerical Kindred	6
Managers, Officials & Proprietors	7
Semi-Professionals	8
Professionals	9

Education	Score
No School	1
Some Elementary	2
Elementary Completed	3
Some High School	4
Completed High School	5
Some College	6
Completed College	7
Some Graduate Work	8
Completed Graduate Work	9

In describing the scale, Meier and Bell state that the differentiation of socioeconomic status,

> "...was determined by a composite score based on a simple average of scores given for each respondent's occupation, education and income...The range on the composite index of socioeconomic status is 24, with a lowest possible score of 3 and a highest of 27. Low SES was dichotomized by grouping together persons with scores from 3 to 17 and high SES by

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

grouping those persons with scores from 18 to 27."[14] Although the Meier-Bell scale dichotomizes individual social status on the basis of income, occupation, and educational variables, it was felt that important social status factors also related to residential indices. Accordingly a residential scale was devised by the author combining the type of dwelling, payment for housing and location of dwelling. This was in keeping with studies involving similar variables by Kahl[15] and Yee[16].

TABLE II

Residence Scale of SES

Type of Dwelling	Score	
Public Housing/Projects	1	Low SES
Apartment	2	
Flat	3	Transition
Rent Home	4	Mid/High SES
Own Home	5	

Housing Payment	Score	
$0-$75	1	
$76-$100	2	Low SES
$101-$150	3	
$151-$200	4	
$201-$275	5	
$276-$325	6	Mid/High SES
Over $325	7	

Neighborhood/Locale	Score	
Projects/Public Housing	1	Low SES
Predominantly Black	2	
Integrated	3	Mid High
Predominantly Caucasian	4	

[14]Meier and Bell, Ibid., p. 202.

[15]Kahl, Op cit., p. 323.

[16]Yee, Op. cit., pp. 33-35.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The range of score values was 13 with the lowest value being a score of 3 and the highest 16. Low SES in relation to residential variables was defined as score values from 3 to 8, while mid/high residential SES ranged from 9 to 16. A composite SES score was then obtained on each subject by combining the Meier-Bell and Residential SES Scales as follows:

TABLE III

Total SES (Social Class) Rating Scale

Meier-Bell:	Low SES rating: 3-17 High SES rating: 18-27
Residential SES Scale:	Low Residential Status: 3-8
Residential SES Scale:	Mid/High Residential Status: 9-16
Total (Composite) SES Score:	Low Social Class: 6-25 Mid/High Social Class: 27-43

It can be seen that by assigning a point value to the six variables comprising social class status among Black populations (Income, Occupation, Education, Type of Dwelling, Housing Payment, Neighborhood/Locale) a reasonably exact point differentiation can be ascertained. For statistical and research purposes, the two status groupings were dichotomized at the transition level of 26 points; all cases which received this score were discarded. Throughout the study, efforts were made to secure as clear-cut a social class dichotomy as possible. In comparing the composite scores for the Lower Class Ss, the average score was 18.6, while for the Middle Class Ss, the average score was 34.7, thus assuring an adequate separation between social class variables for the groups comprising the study. With reference to the determination of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

class status for each subject, this area will be discussed in greater depth in conjunction with data regarding recruitment of Ss and procedural information in the "Methodology" subsection of this chapter.

The Instruments

BITCH: Description

The Black Intelligence Test of Cultural Homogeneity (BITCH) is a 100-item, culture-specific, multiple choice vocabulary test. The content of the instrument was drawn exclusively from the Black experience, as well as the Dictionary of Afro-American Slang.[17] The original word list consisted of 175 items selected at random, then subjected to item selection and discrimination procedures, so as to eliminate ambiguous and duplicative words. Finally, the words were administered to Black and Caucasian experimental groups to identify relevant criteria and select the more potent discriminator words. A group of judges further culled the items along the above dimensions. In the end, the 100 best of the original 175 items were chosen.[18]

Standardization Sample

The test subjects norming the instrument were 100 Black and 100 Caucasian high school students. Williams states that "half the Ss were from low socioeconomic levels, whereas

---

[17] C. Major, Dictionary of Afro-American Slang, New York, N.Y.: International Publishers, (1970).

[18] Williams, Op cit., p. 6.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the other half came from middle income levels."[19] The means and standard deviations of the scores of the two groups are listed in Table IV:

TABLE IV

Means and Standard Deviations of Blacks and Whites on the BITCH 100[20]

	White	Black	Combined W & B
N	100	100	200
MEANS	51.07	87.07	69.07
SD	16.20	6.97	21.92

Test Administration

The administration of the BITCH requires less than one-half hour time duration. The directions are as follows:

> "Below are some words, terms and expressions taken from the Black experience. Select the correct answers and put a check mark in the space provided on the right of the test sheet. Remember, we want the the correct definition as Black people use the words and expressions. There is no time limit. 20 to 30 minutes should be sufficient to complete the test. Go ahead."[21]

Reported Reliability and Validity Indices

Test-retest reliability coefficients were plotted, as were split-half data. Both measures are seen in Table V.

---

[19]Williams, Ibid.

[20]Williams, Ibid., p. 7.

[21]Williams, Ibid., p. 14.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE V

BITCH 100: Split Half and Test-Retest Reliabilities[22]

	N	MEAN	SD	S-H REL.	TEST-RETEST
Black	100	87.07	6.97	.90	.88
White	100	51.07	16.20	.86	.84

With reference to BITCH validation, Williams, as previously stated, does not obtain measures concurrent with such established instruments as, for example, the Wechsler or Stanford Binet scales. He feels that these tests contain a bias against Black populations and are therefore inappropriate for such comparisons. Indeed, a major thrust of this study is to determine just this point. Williams, however, does report several followup studies which offer some measure of validation. The BITCH was administered to a group of Neighborhood Youth Corps high school "drop outs," where the age range was from 16-18; 28 Ss (17 female, 11 male) participated in the study. Additionally, the California Achievement Test was also given this sample. Interestingly, the two "verbal" CAT subtests (Reading and Language) showed reasonably high correlations with the BITCH. Although Williams fails to note the significance level, his reported score values, if accurate, indicate that the correlation between the BITCH and the Reading subtest of the CAT is significant at the .05 level. This can be

[22]Williams, Ibid., p. 13.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

seen in comparing the correlation coefficient of .39 with the tables listing critical values of r, where with an n of 28 (26df) an r of .374 or better is required to achieve an Alpha level of .05. The table relating to this study follows:

TABLE VI

Coefficients of Correlation Between BITCH and CAT Scores[23]

Test	n	Mean	SD	r with BITCH
Reading	28	7.60	1.96	.39*
Language	28	7.69	1.81	.33
Math	28	7.34	1.34	.18
BITCH	28	80.79	9.20	--

*Significant at .05

Another source of comparison comes from cross-validation data where the BITCH was administered to an interracial sample (25 Black, 13 White) at Tougaloo and Millsaps Colleges, Mississippi in 1972. The students were taking a "methods and statistics" course. Reported means were 71 for the Black students, and 59 for the Caucasians, which although considerably lower for the Black sample than the St. Louis study, still demonstrated the differential between the racial groups. This difference score is a cornerstone of William's culture-specific test approach. An additional source of comparative data was reported in December 1972 from Boston University samples where 19 Caucasian graduate students were administered the BITCH; their

[23]Williams, Ibid., p. 11.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

mean score was 60.[24] This roughly parallels the means for White students in the St. Louis and Mississippi studies, thus giving somewhat greater credence to William's assumptions underlying the BITCH, at least in the reported Black/Caucasian score differentials.

Both the Boston and Mississippi studies (1972) used a non-typical sample (college and graduate students) with very small reported numbers. Although Williams indicates in the BITCH test manual that additional validation studies are underway involving, a "sample of approximately 54,000 Black students in four regions of the country,"[25] to the best knowledge of the author, no followup of this scope has been reported in the literature in the six years since publication of the manual (1972). In Williams' eyes, "the final validation of the BITCH will not rest upon how well it correlates with established ability tests, but how well it works out in practice."[26]

Wechsler Intelligence Scale for Children-Revised (WISC-R)

The Wechsler Intelligence Scale for Children-Revised, is an individually-administered test of general ability which is divided into twelve subtests (including two alternates). The subtests are further grouped according to a Verbal and Performance (primarily non-verbal) scale. Table VII gives the breakdown of these groupings.

---

[24]Williams, Ibid., p. 12.

[25]Williams, Ibid., p. 11

[26]Williams, Ibid., p. 10.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE VII

Verbal and Performance Scale of the WISC-R[27]

Verbal Scale	Performance Scale
1. Information	2. Picture Completion
3. Similarities	4. Picture Arrangement
5. Arithmetic	6. Block Design
7. Vocabulary	8. Object Assembly
9. Comprehension	10. Coding
11. (Digit Span)	12. (Mazes)

The subtests are numbered in the order given; administration of the WISC-R takes approximately two hours. The scale must be given by an examiner qualified and trained in its administration, scoring and interpretation. In the standardization of the WISC-R, all twelve subtests were administered, although only ten are utilized to establish I.Q. scores. The alternate tests are to be used when time permits, another subtest is invalidated, or as a means of supplementing the other subtests in view of the "qualitative and diagnostic information they add."[28]

In the preface to the WISC-R manual, Wechsler states that,

> "...The revised WISC, like the scale it succeeds, has been designed and organized as a test of general intel-

---

[27]Wechsler, Op. cit., p. 8.

[28]Wechsler, Ibid., p. 9.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

> ligence. Its author believes that general intelligence exists; that it is possible to measure it objectively; and that, by so doing, one can obtain a meaningful and useful index of a subject's mental capacity. He also believes that the much challenged and berated I.Q., in spite of its liability to misinterpretation and misuse, is a scientifically sound and useful measure, and for this reason he has retained the I.Q. as an essential aspect of the revised Scale."[29]

Standardization Sample

One of the primary considerations in re-standardizing the WISC-R was in "modification or elimination of items felt by some test users to be ambiguous, obsolete, or differentially unfair to particular groups of children."[30] This resulted in changing many items which are detailed, item by item, and subtest by subtest in the WISC-R manual.[31] In addition to former standardization procedures with respect to total numbers (2,200 children; 200 at each of eleven age levels), Age 6 1/2 through 16 1/2 years), Sex (equally divided throughout the scale), and other demographic variables, racial considerations were carefully delineated. Whites and non-Whites were included in the sample, "in the same proportion found in the 1970 census for the age range tested."[32]

Sample representation very closely parallels population characteristics, as reported by the 1970 census figures. Concerning the racial variable, it should be pointed out that of

---

[29]Wechsler, Ibid., p. iii.

[30]Wechsler, Ibid., p. 10.

[31]Wechsler, Ibid., pp. 12-16.

[32]Wechsler, Ibid., p. 17.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the total number of non-Whites in the standardization sample (N=330), 305 (92.4%) were Black; this sample exceeds the actual proportion of Blacks (91.3%) to non-Whites within the United States.[33] The actual number of Black children who participated in the WISC-R standardization thus represents the significant majority of all non-Whites in the nation (15%), according to the 1970 census figures.[34] Table VIII compares the total WISC-R sample by geographic region and race to the actual percentages extant in the United States' population.[35]

TABLE VIII

Standardization Sample by Geographic Region and Race as Compared to the 1970 U. S. Population Characteristics[36]

	Northeast			Northcentral			South		
	Wh.	N.W.	Tl.	Wh.	N.W.	Tl.	Wh.	N.W.	Tl.
Total Sample (N=2200)	19.0	2.8	21.8	26.1	3.0	29.1	24.0	7.6	31.6
% In U.S. Pop.	29.3	2.6	22.9	25.7	2.9	28.6	23.7	7.6	31.3
	West			All Regions					
Total Sample	15.9	1.6	17.5	85.	15.0	100.00			
% In U.S.P.	15.3	1.9	17.2	85.	15.0	100.00			

[33]Wechsler, Ibid., p. 19.

[34]U.S. Bureau of the Census, Census of Population: 1970 General Population Characteristics, Final Report P.C.(1)Bl. U.S. Summary, U.S.Govt. Printing Office, Wash.D.C., 1972, Table 56.

[35]Wechsler, Op. cit., p. 20.

[36]U.S. Bureau of the Census, Ibid.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reliability and Validity Data

Reliability coefficients are split-half correlations corrected by the Spearman-Brown formula, with the exceptions of the Digit Span and Coding subtests which utilized test-retest procedures. Verbal, Performance, and Full Scale I.Q.'s, "have high reliabilities across the entire age range, the average coefficients being .94, .90, and .96 respectively."[37] Standard errors of measurement (SE_m), represent the error fluctuation around any particular score and an indicator of the confidence one can place in making judgments about an individual's true ability on any particular test instrument. In the case of the WISC-R, the average SE_m values for all age ranges tested were 3.60 for Verbal I.Q.'s, 4.66 for Performance I.Q.'s, and 3.19 for Full Scale I.Q.'s.[38]

Concurrent validation was investigated between the WISC-R and other measures of intelligence, specifically the Wechsler Preschool and Primary Scale of Intelligence (WPPSI), the Wechsler Adult Intelligence Scale (WAIS), and the Stanford-Binet Intelligence Scale (Form L-M, 1972 norms). Correlational measures reveal coefficients of .82 with the WPPSI, .95 with the WAIS, and .73 with the Stanford-Binet.[39] With regard to the latter statistic, this value is similar to those

---

[37]Wechsler, Op. cit., p. 27.

[38]Wechsler, Loc. cit.

[39]Wechsler, Ibid., p. 48-51.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

derived from comparisons between the 1949 WISC and the S-B (Form L-M, 1960 norms).[40] Moreover, Hartlage and Steele (1977) in comparing WISC-R scores of a primarily Black sample (81%) of first and second graders with obtained achievement scores (and actual grades), found correlations which were significant at the .01 level or better when measured against indices of reading, spelling, writing, and arithmetic.[41] Score differences between WISC-R Verbal and Performance measures for a Black sample (N=305) showed no significant differences in a study by Kaufman and Doppelt (1976), although there was an interracial difference of about one standard deviation in favor of Caucasian groups.[42] The absolute I.Q. values were narrowed when score differences between occupational categories were obtained; this difference held true for Black as well as White groups and at all age levels.[43] Thus to the extent that parent's occupational group (e.g., "professional" vs. "blue collar"), taken by itself, correlates with indices of social class, WISC-R I.Q.'s "are clearly related to the child's socioeconomic level."[44]

---

[40]Wechsler, Ibid.

[41]L.C. Hartlage and C.T. Steele, "WISC and WISC-R Correlates of Academic Achievement, "Psychology in the Schools, V. 14, (1977), pp. 15-18.

[42]Alan Kaufman and Jerome Doppelt, "Analysis of WISC-R Standardization Data in Terms of Stratification Variables," Child Development, V. 47, (1976), pp. 165-171.

[43]Kaufman and Doppelt, Ibid., p. 168.

[44]Kaufman and Doppelt, Ibid., p. 169.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

California Achievement Test

The California Achievement Test (CAT), 1970 edition,

> "...was designed for the measurement, evlauation, and analysis of school achievement. The emphasis is upon content and objectives in the basic curricular areas of reading, mathematics, and language. The intended measurement is one of performance in these curricular areas...a major aim of the 1970 revision of the CAT is...to provide new norms and derived scores based on a truly national sample which will aid users in making more meaningful interpretation of test results."[45]

The CAT is a multiple-choice, group-administered, test of achievement with sections pertaining to skills in reading (vocabulary and comprehension), mathematics (computation, concepts and problems) and language (capitalization, punctuation, usage, and spelling). As previously mentioned, only the reading portion of the CAT was used in this study, due to the lower correlations obtained in an earlier study by Williams (1972) between the BITCH and the mathematics/language portion of the CAT.[46]

The Vocabulary subtest of the CAT (reading) is a timed (ten minutes) section containing forty items each; a stimulus word, in context, is presented along with a list of four alternatives. The student must choose the best alternative.

The Comprehension subtest (timed 40 minutes) contains two parts, with the first presenting six items designed to assess the student's knowledge of reference skills, while the

---

[45]Ernest Tiegs and Willis Clark, California Achievement Tests, Examiners Manual, 1970 Ed., Monterey, CA.: CTB, McGraw Hill, 1970.

[46]Williams, Op. cit., p. 11.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

second section presents five reading passages relating to materials found in the general school curriculum. The student's comprehension is evaluated by means of his supplying of answers to thirty-nine questions designed to test, "relationships, inferences, recall of facts, and identification of main ideas (of the reading passage)."[47]

Standardization Design and Procedure

Standardization testing was accomplished during the early months of 1970 on a nationwide sample of 203,684 students. This sample was gathered in a random, stratified design with the variables of geographic region, type of school district, type of community, and grade level being considered. Each level of the test was, "standardized at the grades for which they were to be used; adjacent levels were administered to the same grade where they overlapped."[48]

Score Reportage and Concurrent Validation

All raw scores derived from the CAT can be converted to grade level measures, National and Local Percentile rankings, Obtained Grade Equivalent scores, Achievement Development Scale scores, National or Local Stanines. Extensive Norms tables exist in the Examiner's Manual for most of the above-named score conversions.[49]

---

[47]Tiegs and Clark, Op. cit., p. 7.

[48]Tiegs and Clark, Op. cit., p. 7

[49]Tiegs and Clark, Ibid., pp. 49-77.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

With reference to measures of concurrent validation (particularly with WISC and Stanford-Binet comparisons), Jenkin, et al, (1964),[50] Nelson and Hudson (1969)[51] Washington and Teska (1970),[52] all report significant correlations at the .01 level between the WISC, Standord-Binet, and the CAT.

The Method

Since a "normal" (i.e., school-attending, grade-appropriate) sample was desired, original efforts related to obtaining the sample were directed at school counselors and administrators in the various school sites within the San Francisco Bay Area. After preliminary telephone contacts with key personnel in each "target" school, a meeting was arranged where the counselors, principal, pupil service workers and others were in attendance. At this time, the purpose of the study was explained in as much detail as necessary to clarify its nature, importance, and scope. All questions were answered; many of those more frequently asked related to time allotments, necessary scheduling, and informing of both the student and his parents. One individual, usually a counselor, was chosen as

---

[50] N. Jenkin, G. Spivach, M. Levine, W. Savage, "Wechsler Profiles and Academic Achievement in Emotionally Distrubed Boys," Journal of Consulting Psychology, V. 28, (1964), p. 290.

[51] C.M. Nelson and F. G. Hudson, Prediciting the Reading Achievement of Junior High School EMR Children,"American Journal of Mental Def., V. 74, (1969), pp. 415-420.

[52] E.D. Washington and J.A. Teska, "Relations Between the WRAT, CAT, Stanford-Binet, and Ill. Test of Psycholinguistic Ability," Psychological Reports, V. 26, (1970), pp. 291-294.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

as the chief liaison individual at that particular school site. His/her responsibility was to provide an alphabetical list of the names of all likely subjects, meet with them, briefly explain the study in a general and non-specific manner, then solicit their cooperation and participation. If each individual agreed, a letter of permission was given them to take home and get signed by their parents. Along with the permission letter, the parent was requested to call the contact worker if any questions were forthcoming. The permission letter is indicated below:

Dear Parent/Guardian:

Your cooperation is solicited and your permission requested so that your son/daughter may participate in a study involving appropriate applications of testing with minority groups. The maximum amount of time involved would be between three to four hours duration. All results would be held in confidence; no names would ever be mentioned. Following the study, your child's individual performance could, if requested, be discussed confidentially with you, the parent or guardian.

This is useful, important and necessary research designed to assist minority children in their future work, both in school and later life; we hope you can further this significant research study.

I hereby give permission for my son/daughter to participate in the minority research study. I understand that all results will be held in complete confidence with respect to the identity of the participants.

_____
Parent/Guardian

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Following the return of the permission letter, the on-site contact person then arranged with administrators and teachers for the necessary time release from class, testing rooms, schedule coordination, and other administrative details. In effect a "pool" of subjects was created which fit the variables considered important in the study as to age, race, sex, grade, average achievement in school, and social class. With respect to the latter variable, and prior to selection of the individual for actual participation in the study, the on-site liaison was given a sheet of questions to ask the student and his parents, so as to determine the indices of social class; this sheet is reproduced below:

Determination of Social Class Status

Questions to be asked of participants and parents.

Category I. Residence Location & Type of Family Housing

1. What is your address?
2. How would you describe your neighborhood?
   a. Predominantly Caucasian (White)
   b. Integrated
   c. Predominantly Black
   d. Projects/public housing
3. In what type of housing do you live?
   a. Owned home
   b. Rented home
   c. Flat
   d. Apartment
   e. Public housing
4. Which of the following categories would describe your monthly payment for housing?
   a. $0-$75
   b. $76-$100
   c. $101-$150
   d. $151-$200
   e. $201-$275
   f. $276-$325
   g. Over $325

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Category II. Occupation of Head of Household

1. What sort of work does your father/mother/guardian/yourself do? (Note: If respondent indicates "no work" then question further to ascertain occupational status (e.g., retired, welfare/AFDC, etc.).
2. If both parents work what does he/she/yourself do?

Category III. Education of Head of Household

1. (Asked of respondent: Parent, guardian, and/or study participant.) How much education (schooling) have you (parent) had (i.e., how far have you gone in school)?
   a. No school
   b. Some elementary
   c. Elementary completed
   d. Some high school
   e. Completed high school
   f. Some college
   g. Some graduate work
   h. Completed graduate work

Category IV. Income of Head of Household

1. In general, can you give me some idea of the family income? (Note: If both parents work, attempt some breakdown of income as per "primary" vs. "secondary" job/income.)

Upon the return of all necessary information, the potential subject was placed in the test "pool" for possible inclusion in the study; no subject was selected unless all necessary permission and data forms were completed, and the individual was willing to proceed. From this "pool" of Ss, every fifth individual was selected. This tended to further randomize the selection of individuals participating in the study.

Testing Procedure

All participants in the study were administered the BITCH. CAT-R, and the WISC-R; all instruments were given in randomly counterbalanced order to control for test effects. Ss were given both the CAT and the BITCH in a small group, while the WISC-R was given individually to each participant. Once an

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

individual was started on a test sequence, the entire battery of three tests was given within a five day period; hence as little time delay as possible between testing sessions existed. The entire WISC-R was administered; although testing time was increased slightly, it was felt that the additional data obtained by the performance (non-verbal) portion of the WISC-R might prove useful in analyzing interaction effects between the BITCH and non-verbal intelligence indices.

The examiners used in this study were all trained and California certificated School Psychologists; all examiners were familiar with the test instruments used in the study, as well as being experienced with the practices of testing, whether individually or in samll-group presentations. All examiners held the Masters degree in test-related fields of study. Of the eleven examiners employed in the testing procedures, four were male and seven female. From a racial standpoint, eight were Caucasian while three were Black. It was felt that the use of an interracial, non-sex-differentiated group of qualified psychological examiners served to increase the degree of generalization one might draw from the data, especially since both male and female, as well as Black and Caucasian examiners were involved; although numerous recent studies have shown the race of the examiner and the race of the subject do not prove to be a significant variable (Barnebey, 1973; Caldwell & Knight, 1970; Dill, 1972; Hall, Reder & Cole, 1975; Samuel, et al., 1976; and Scott, et al., 1976).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

An observation of the impromptu comments of a number of subjects should be noted in passing. Many of them remarked on the fact that a considerable number of BITCH items were unfamiliar to them and they were merely "guessing" at the answers. Others indicated that the meanings of many of the terms had changed (i.e., the "correct" choice was not among the four possibilities listed). Still others asked in a rather sarcastic manner if, for example, "a knowledge of what hominy grits were" was really an important factor? In general, most of the participants were neutral to positive in manner regarding the WISC-R, but were at best neutral in attitude with respect to the BITCH.

Procedural Followup

Following complete test administration, each subject was thanked for his/her participation. Test scores were reduced to a composite face sheet which also included variables relating to social class and identifying data. The Individual Subject Face Sheet is reproduced below:

Test Comparison Face Sheet

Name: Address:
Sex: School:
Birthdate: Grade:
Age:

Social Class Data:
Parents' Income: Single: Both:
Welfare: SSI: Other:
Parents' Occupation:

Parents' Education (highest level reached):

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Parents' Residence:
Own Home___
Rent Home___
Flat___
Apt.___
Pub. Hsng.___

Mo. Pmt.
Housing
$_____

Locale
Pred. Cauc___
Integrated___
Pred. Black___
Proj./Pub.Hsng.___

Total Numerical SC Rating_____
MC:_____ LC:_____

TEST DATA	WISC-R:	FSIQ_____	CAT-Reading
BITCH	VIQ_____	PIQ_____	Vocab.
RS:_____	Info._____	P.C._____	RS_____
	Sim._____	P.Ar._____	Gr. Lev._____
	Arith.____	Bl.Des.___	
	Vocab.____	Cod._____	Comp.
	Comp._____		RS_____
			Gr. Lev._____
			Tot. Read.___

Comments:_____

_____

Additional Information

After the test scoring procedures and reduction of all pertinent information to the individual's face sheet were accomplished, division of the test groups into two dichotomies, based upon the numerical SES rating (as described earlier) and the sex of the participant was accomplished. These groups were identified as MM (Male; Middle Class), ML (Male; Lower Class), FM (Female; Middle Class), and FL (Female; Lower Class). Within this grouping each individual was assigned an identification number which replaced all identifying information (e.g., name, address, school, birthdate), so as to preserve the Ss anonymity. The single copy of the list matching name and ID number is kept by the author and will not be released. This is not only for the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

subject's protection, but also to fulfill the terms under which test subjects were originally recruited. At a later date, the author will meet with the families of those participants who requested a followup contact for further explanation of the study and the individual performance of their son or daughter. Even this contact will be of a general nature, so that sensitive (e.g., I.Q. scores) information which might compromise the participant, will be carefully screened. This is to insure the general anonymity and individual protection of every subject who participated in this study.

Chapter Summary

This chapter has addressed itself to a description in depth of the project sample, the instruments employed in testing the subjects, and the methods which were utilized gathering the data in the study. Within each chapter subheading, all supporting data were also included, so as to acquaint the reader with all procedures and operations in this research project.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER IV

STATISTICAL ANALYSIS

Overview

This chapter will concern itself with an analysis in depth of the data obtained from Ss scores in the study. After detailing preliminary considerations as to relevant and related general information, restatement of hypotheses, type of statistic employed and rationale, the reporting of measures of central tendency, variability and other descriptive data will be forthcoming. Finally, measures of inference, including t-tests, Analyses of Variance, and appropriate correlational techniques will be discussed in relation to the hypotheses explored in this research design.

Preliminary Considerations

All computations were performed at the Stanford Center for Information Processing, Stanford University, Palo Alto, California on March 27 and 28, 1978. The computer utilized in analyzing results was the IBM 370/168, while the "software" package which programmed the data was the Statistical Package for the Social Sciences (SPSS).[1] After dichotomizing the sample into four groups based upon sex, social class, and

---

[1]Norman Nie, C. Hadlai Hull, Jean G. Jenkins, Karin Steinbrenner, Dale H. Bent, Statistical Package for the Social Sciences (2nd Ed.) Manual, New York, N.Y.: McGraw-Hill (1975).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

assignment of an identifying number for each individual within his or her group, all data was coded for keypunching prior to the actual computer run. The maximum number of variables accorded each subject on the basis of sex, age, grade, social class and test scores were twenty-two. The breakdown of coding data for two Ss in the ML group are listed in Table IX.

TABLE IX

SPSS Coding Data (Legend)

ID #	Sex	Age	Grade	NumSC	SC	BITCH	FSIQ	VIQ	Info	Sim	Arith	Vocab	Comp	PIQ	PComp	PArr	BlkD	Codng
1	1	14	9	24	1	44	114	115	14	12	12	11	14	109	14	8	13	11
2	1	14	9	12	1	52	112	108	10	14	11	11	11	114	14	10	12	12

Sex: 1=Male; 2=Female Soc. Class: 1=Lower; 2=Middle

CAT Vocab	CAT Read	CAT Total
12.4	13.3	12.9
4.8	8.2	6.5

The total number of Ss in each group were as follows: Male Lower Class (ML), N=16; Male Middle Class (MM), N=15; Female Lower Class (FL), N=18; Female Middle Class (FM), N=16. Thus the total number of participants were 65, with 31 male and 34 female. The SES breakdown consisted of 34 lower class and 31 middle class. Each of the "lines" of coded data were key-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

punched onto IBM cards for computer analysis. The computer was also instructed (by means of "save file" cards) to retrieve all cards and data for followup analysis or re-programming at a later date should additional information (or future studies) so warrant.

Restatement of Hypotheses

Statistical:	(H_0)I:	There is no relationship between the BITCH and the Verbal scale of the WISC-R.
Alternate:	(H_1)I:	There will be a significant relationship between the BITCH and the Verbal scale of the WISC-R.
Statistical:	(H_0)II:	There is no relationship between the BITCH and Verbal Achievement as measured by a conventional achievement test (California Achievement Test).
Alternate:	(H_1)II:	There will be a significant relationship between the BITCH and Verbal Achievement as measured by the California Achievement Test.
Statistical:	(H_0)III:	There is no relationship in scores on the BITCH between a middle and lower class Black sample.
Alternate:	(H_1)III:	There will be a significant relationship in scores on the BITCH between a middle and lower class Black sample, with low social class Blacks scoring higher than middle social class Ss.

Statistical Method and Rationale

Hypotheses I and II involve comparisons of BITCH scores with Verbal I.Q. or achievement scores. As such, the data lend themselves to analysis by means of correlational techniques. This is consistent with parallel studies involving other test instruments, where similar statistics have been

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

utilized.[2] Sattler[3] and Anastasi[4] report an extensive number of research designs involving concurrent validation; all such studies employ correlational methods in their analysis. The data obtained from Ss test scores which relate to Hypotheses I and II are continuous (as contrasted with dichotomous) sets of variables. Moreover, all observations are independent (i.e, the selection of any particular case for inclusion in the study did not bias the chances of any other case for consideration). Both intelligence and achievement variables are normally distributed and lend themselves to interval scale measurements, so that the Pearson Product-Moment correlational technique is the treatment of choice in this instance. With regard to sample size, an N of 65 with 63 df requires an r of .25 or higher to achieve significance for a two-tailed test.[5]

Although the assumptions underlying the first and second hypotheses comply with the choice of a parametric statistic, Hypothesis III, dealing with scoial class indices, assumes only ordinal scale data for one set of variables. The dichotomizing of social class into a "lower" and "middle" group assigns a "rank" or order which does not on the face of it, lend itself to parametric statistics. Disagreements exist

---

[2]D.J. Reschly & J.E Reschly, "Validity of WISC-R Scores in Predicting Achievement and Attention for Four Sociocultural Groups," Unpublished Manuscript, Iowa State University (1977).

[3]Sattler, Op. cit., pp. 425-429.

[4]Anastasi, Op. cit., pp. 327-357.

[5]Hardyck & Petrinovich, Op. cit., p. 308.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

in the statistical literature regarding the need to use only non-parametric measures on data that are less than interval scale measurements. Hardyck and Petrinovich state,

> "In general, we are perfectly safe in calculating any statistic we want on any set of measurements that have the properties of an ordinal scale. There is definitive evidence that statistics calculated on ordinal measurements are just as reliable and meaningful as statistics calculated on interval or ratio scales of measurements."[6,7]

Although the primary statistic employed in analyzing Hypothesis III will be non-parametric in nature (Kendall Rank Correlation Coefficient) it was decided to run H_o III variables using both Kendall's "tau" correlation, as well as the Pearson r to ascertain any differences. Regarding the comparison of Kendall "tau" to Pearson r, Siegel indicates,

> "When used on data to which the Pearson r is properly applicable, tau...has efficiency of 91 percent. That is tau is approximately as sensitive a test of the existence of association between two variables in a bivariate normal population with a sample of 100 cases as is the Pearson r with 91 cases."[8]

In addition to the correlational measures previously indicated, it was decided to perform an Analysis of Variance between all test scores and social class indices in order to ascertain any interaction effects, as well as the significance of this variable relative to intelligence and achievement.

---

[6]Hardyck & Petrinovich, Ibid., p. 27.

[7]B.O. Baker, C.D. Hardyck & L.F. Petrinovich, "Weak Measurements vs. Strong Statistics: An Empirical Critique of S.S. Stevens Proscriptions on Statistics," Educational and Psychological Measurement, V. 26, (1966), pp. 291-309

[8]Sidney Siegel, Nonparametric Statistics, New York, N.Y.: McGraw Hill, (1956), p. 223.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

A t-test was also performed, so as to uncover significant differences, if any, between male and female participants in this study. In all cases, an Alpha level of .05 or better will be considered significant.

Descriptive Statistics

The median age of the subjects participating in the study was 14.7 with a range of 3.0 between the youngest and oldest individuals (13 to 16 years). Average grade level was 9.5, with a range of 4.0 between the seventh and eleventh graders. Ninth graders comprised 40.0 percent of the sample; tenth graders 33.8 percent, eleventh graders 16.9 percent, seventh graders 4.9 percent, eighth graders 4.9 percent. Thus the majority of Ss taking part in the study (73.8 percent) were in the ninth and tenth grades.

Social Class: Two approaches to assessing effects of social class status were considered; the first involved assignment of all individuals to a social class category. Thus all lower class Ss were given a rating of one, while middle class Ss were assigned a rating of two. This division was based upon their previously-determined numerical SES standing, but was labeled "categorical social class." The second approach to dichotomizing this variable was directly related to the assignment of a numerical rating to each individual based upon indices of social class discussed in an earlier section of this paper. A rating of 6 to 25 was considered to be low SES, while ratings of 27 to 43 were considered middle SES; this grouping was labeled "numerical social class."

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Categorical social class breakdowns indicate 34 Ss (52.3 percent) in the first (LC) group and 31 Ss (47.7 percent) in the second (MC) ranking. Numerical social class means were 26.3 (precisely at the divisional break between the two groups), with a standard deviation of 9.2. Percentages indicate 52 percent of the sample to be numerical lower SES and 49 percent to be numerical middle SES. The range was 32.0 between a low point rating of 11 and a high of 43. Means of LC participants were 18.6, while means for MC subjects were 34.7.

BITCH Results: The mean score for all Ss taking the BITCH was 61.9 with an SD of 11.5; the range was 44.0 between low scores of 39 and a high of 83. By Categorical SES, lower class means were 60.6 (SD 10.9) and middle class means were 63.3 (SD 12.2). All scores, frequencies, adjusted and cumulative percentages are reported in the following table.

TABLE X

Scores, Frequencies, Adjusted and Cumulative Percentages of all Ss on the BITCH

Score	Freq.	Adj.%	Cum.%	Score	Freq.	Adj.%	Cum.%
39	1	2	2	60	3	5	46
41	1	2	3	61	1	2	48
43	1	2	5	62	2	3	51
44	1	2	6	63	3	5	55
45	2	3	9	64	3	5	60
46	1	2	11	66	2	3	63
47	1	2	12	69	4	6	69
49	1	2	14	71	3	5	74
50	1	2	15	72	3	5	78
51	4	6	22	73	2	3	82
52	3	5	26	75	4	6	88
53	4	6	32	77	1	2	89
54	1	2	34	78	2	3	92
55	2	3	37	80	3	5	97
57	2	3	40	81	1	2	98
58	1	2	42	83	1	2	100

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

WISC-R Results: Cumulative means (102.3) and standard derivations (16.1) on the WISC-R closely parallel national norms (M=100; SD=15). Mean Verbal I.Q. for the entire group was 102.5 (SD=16.8), and Performance I.Q. means were 101.6 (SD=15.6). Comparison of WISC-R I.Q.'s by social class, however, reveals an across-the-board difference in means of about one standard deviation. Table XI illustrates the score differentials.

TABLE XI

Score Differentials by Social Class on the WISC-R

Social Class	Verbal I.Q.		Performance I.Q.		Full Scale I.Q.	
	Mean	SD	Mean	SD	Mean	SD
Middle SC	109.5	13.7	109.2	14.4	110.5	12.7
Low SC	96.0	17.0	94.6	13.5	94.9	15.4

Similar results are seen when I.Q. subtest scores are broken down by social class as the following table reveals.

TABLE XII

Subscale Mean-Differentials on the WISC-R by Social Class

Soc. Class	Verbal					Performance			
	Inf.	Sim.	Ari.	Voc.	Com.	P.C.	P.A.	B.D.	Cod.
Middle	10.4	12.3	10.8	11.4	13.1	11.5	11.6	11.0	12.2
Low	7.6	9.1	9.0	9.1	10.7	9.6	10.1	8.2	9.5
All	9.1	11.0	9.8	10.2	11.9	10.5	10.8	9.5	10.8

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

It should be noted that WISC-R subscale means and S.D.'s for a statistically "average" youngster should yield an adjusted expectancy score of ten (10) with a standard deviation of three (3).[9] This "ideal" score should be kept in mind when consulting the preceding table.

CAT Results: The median grade level for the entire group of Ss was 8.0 with a mean grade of 8.3 and a standard deviation of 3.2. The range was 12.2 between low scores of 1.4 and highs of 13.6. A breakdown of total reading scores into their respective components of Vocabulary and Comprehension yields similar results, across social class lines, as seen in Table XIII.

TABLE XIII

Grade Score Differentials by Social Class on the CAT-Reading Test

Social Class	CAT-Vocab.	CAT-Comp.	CAT-Tot.Read.
Middle SC	10.2	9.6	9.9
Low SC	7.2	6.5	6.8
All SC	8.7	8.0	8.3

Inferential Data

Pearson product-moment coefficients of correlation were run on twenty sets of variables whose means and standard deviations appear in the following table.

[9] Wechsler, Op. cit., pp. 49-50.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE XIV

Means and Standard Deviations on all Correlated Variables

Variable	Cases	Mean	Std. Dev.
Age	65	14.72	0.91
Grade	65	9.53	0.99
Num SC	65	26.30	9.19
SC	65	1.48	0.50
BITCH	65	61.89	11.59
Full I.Q.	65	102.34	16.08
Verbal I.Q.	65	102.46	16.85
Info	65	9.06	3.28
Sim	65	10.95	3.71
Arith	65	9.82	2.59
Vocab	65	10.22	3.10
Comp	65	11.86	3.84
Perf. I.Q.	65	101.55	15.65
P. Comp.	65	10.46	3.47
P. Arr.	65	10.82	3.01
Blk. D.	65	9.51	2.75
Code	65	10.78	3.68
CAT Vocab.	65	8.66	3.29
CAT Comp.	65	8.01	3.27
CAT All	65	8.28	3.20

The obtained Pearson Coefficients of Correlation were arranged into a correlation matrix. These coefficients are presented in Table XV. A non-parametric correlational technique (Kendall's Tau) was also run on all variables; results are presented in Table XVI. An Analysis of Variance by categorical social class was performed on all variables within the study. It should be noted that since Hypothesis III implies directionality, one-tailed tests of significance are employed when comparing means between lower and middle social class groups. A breakdown of pertinent data obtained from the Analyses of Variance and their significance levels appear in Table XVII.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In addition, A T-Test was computed to determine significant differences, it any, between means for males and females. Data relating to this statistic are presented in Table XVIII.

Summary

This chapter has focused upon a statistical analysis of all scores obtained from the participants in the study. Hypotheses and statistical rationale were presented, along with the specification of the descriptive and inferential procedures which organize, quantify, and assign probabilities to the data yielded by the results. A discussion in depth of the nature of the results and their implication will follow in the next chapter.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE XV

Intercorrelations of Variables and Significance Levels for All Ss (N=65)

	AGE	GRADE	NUM SC	SC	BITCH	FULL IQ	VER.IQ	INFO	SIM	ARITH
AGE		.87 S=.001*	-.02 S=.430	.01 S=.437	.29 S=.010*	-.01 S=.457	-.08 S=.268	-.14 S=.141	.02 S=.439	.02 S=.465
GRADE	.87 S=.001*		.20 S=.059	.17 S=.092	.34 S=.003*	.17 S=.088	.11 S=.190	.06 S=.311	.12 S=.166	.17 S=.090
NUM SC	-.02 S=.430	.20 S=.059		.87 S=.001*	.23 S=.035*	.57 S=.001*	.53 S=.001*	.55 S=.001*	.42 S=.001*	.39 S=.001*
SC	.02 S=.437	.17 S=.092	.87 S=.001*		.12 S=.173	.49 S=.001*	.40 S=.001*	.39 S=.001*	.36 S=.002*	.36 S=.002*
BITCH	.29 S=.010*	.34 S=.003*	.23 S=.035*	.12 S=.173		.22 S=.041*	.26 S=.017*	.24 S=.028*	.22 S=.039*	.23 S=.030*
FULL IQ	-.01 S=.457	.17 S=.088	.57 S=.001*	.49 S=.001*	.22 S=.041*		.91 S=.001*	.74 S=.001*	.83 S=.001*	.70 S=.001*
VERBAL IQ	-.08 S=.268	.11 S=.190	.53 S=.001*	.40 S=.001*	.26 S=.017*	.91 S=.001*		.86 S=.001*	.86 S=.001*	.70 S=.001*
INFO	-.14 S=.141	.06 S=.311	.55 S=.001*	.39 S=.001*	.24 S=.028	.74 S=.001*	.86 S=.001*		.66 S=.001*	.53 S=.001*
SIM	.02 S=.439	.12 S=.166	.42 S=.001*	.36 S=.002*	.22 S=.039	.83 S=.001*	.86 S=.001*	.66 S=.001*		.52 S=.001*
ARITH	.01 S=.465	.17 S=.090	.39 S=.001*	.36 S=.002*	.23 S=.030*	.70 S=.001*	.70 S=.001*	.53 S=.001*	.52 S=.001*	
VOCAB	-.18 S=.078	.05 S=.336	.54 S=.001*	.37 S=.001*	.34 S=.003*	.69 S=.001*	.82 S=.001*	.81 S=.001*	.58 S=.001*	.51 S=.001*

*Significant at .05 or better

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE XV (CONTINUED)

Intercorrelations of Variables and Significance Levels for All Ss (N=65)

	AGE	GRADE	NUM SC	SC	BITCH	FULL IQ	VER IQ	INFO	SIM	ARITH
COMP	.02 S=.451	.10 S=.218	.44 S=.001*	.32 S=.005*	.17 S=.083	.82 S=.001*	.90 S=.001*	.68 S=.001*	.76 S=.001*	.58 S=.001*
PER I.Q.	.05 S=.345	.18 S=.077	.49 S=.001*	.47 S=.001*	.10 S=.206	.87 S=.001*	.58 S=.001*	.42 S=.001*	.59 S=.001*	.51 S=.001*
P COM	.26 S=.017*	.38 S=.001*	.32 S=.004*	.27 S=.013*	.03 S=.418	.60 S=.001*	.42 S=.001*	.32 S=.005*	.43 S=.001*	.39 S=.001*
P ARR	-.04 S=.388	.01 S=.459	.28 S=.011*	.24 S=.025*	.21 S=.047	.58 S=.001*	.48 S=.001*	.37 S=.001*	.47 S=.001*	.26 S=.018*
BLK D	.01 S=.478	.13 S=.144	.51 S=.001*	.49 S=.001*	.13 S=.143*	.79 S=.001*	.60 S=.001*	.45 S=.001*	.58 S=.001*	.52 S=.001*
CODE	.01 S=.454	.10 S=.211	.37 S=.001*	.37 S=.001*	.10 S=.206	.37 S=.001*	.14 S=.128	.16 S=.101	.15 S=.110	.20 S=.059
CAT VOC	.04 S=.372	.21 S=.043	.63 S=.001*	.46 S=.001*	.47 S=.001*	.60 S=.001*	.66 S=.001*	.73 S=.001*	.43 S=.001*	.36 S=.001*
CAT COMP	.08 S=.270	.25 S=.024*	.62 S=.001*	.49 S=.001*	.35 S=.002*	.67 S=.001*	.75 S=.001*	.75 S=.001*	.57 S=.001*	.41 S=.001*
CAT ALL	.06 S=.307	.24 S=.029*	.64 S=.001*	.48 S=.001*	.44 S=.001*	.66 S=.001*	.73 S=.001*	.78 S=.001*	.52 S=.001*	.40 S=.001*

*Significant at .05 or better

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE XV (CONTINUED)

Intercorrelations of Variables and Significance Levels for All Ss (N=65)

	VOCAB	COMP	PER IQ	P COM	P ARR	BLK D	CODE	CAT V	CAT C	CAT ALL
AGE	-.18 S=.078	-.02 S=.451	.05 S=.345	.26 S=.017	-.04 S=.388	.01 S=.478	.01 S=.454	.04 S=.372	.08 S=.270	.06 S=.307
GRADE	.05 S=.336	.10 S=.218	.18 S=.077	.38 S=.001*	.01 S=.459	.13 S=.144	.10 S=.211	.21 S=.043*	.25 S=.024*	.24 S=029*
NUM SC	.54 S=.001*	.44 S=.001*	.49 S=.001*	.32 S=.004*	.28 S=.011*	.51 S=.001*	.37 S=.001*	.63 S=.001*	.62 S=.001*	.64 S=.001*
SC	.37 S=.001*	.32 S=.005*	.47 S=.001*	.27 S=.013*	.24 S=.025*	.49 S=.001*	.37 S=.001*	.46 S-.001*	.49 S=.001*	.48 S=.001*
BITCH	.34 S=.003*	.17 S=.083	.10 S=.206	.03 S=.418	.21 S=.047*	.13 S=.143	.10 S=.206	.47 S=.001*	.35 S=.002*	.44 S=.001*
FULL IQ	.69 S=.001*	.82 S=.001*	.87 S=.001*	.60 S=.001*	.58 S=.001*	.79 S=.001*	.37 S=.001*	.60 S=.001*	.67 S=.001*	.66 S=.001*
VERB IQ	.82 S=.001*	.90 S=.001*	.58 S=.001*	.42 S=.001*	.48 S=.001*	.60 S=.001*	.14 S=.128	.66 S=.001*	.75 S=.001*	.73 S=.001*
INFO	.81 S=.001*	.68 S=.001*	.42 S=.001*	.32 S=.005*	.37 S=.001*	.45 S=.001*	.16 S=.001*	.73 S=.001*	.75 S=.001*	.78 S=.001*
SIM	.58 S=.001*	.76 S=.001*	.59 S=.001*	.43 S=.001*	.47 S=.001*	.58 S=.001*	.15 S=.110	.43 S=.001*	.57 S=.001*	.52 S=.001*
ARITH	.51 S=.001*	.58 S=.001*	.51 S=.001*	.39 S=.001*	.26 S=.018*	.52 S=.001*	.20 S=.059	.36 S=.001*	.41 S=.001*	.40 S=.001*
VOCAB		.64 S=.001*	.36 S=.002*	.23 S=.031*	.39 S=.001*	.42 S=.001*	.04 S=.385	.77 S=.001*	.77 S=.001*	.80 S=.001*

* Significant at .05 or better

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE XV (CONTINUED)

Intercorrelations of Variables and Significance Levels for All Ss (N=65)

	VOCAB	COMP	PER IQ	P COM	P ARR	BLK D	CODE	CAT V	CAT C	CAT ALL
COMP	.64 S=.001*		.54 S=.001*	.40 S=.001*	.44 S=.001*	.57 S=.001*	.14 S=.134	.57 S=.001*	.65 S=.001*	.64 S=.001*
PER IQ	.36 S=.002*	.54 S=.001*		.65 S=.001*	.59 S=.001*	.83 S=.001*	.56 S=.001*	.38 S=.001*	.42 S=.001*	.42 S=.001*
P COM	.23 S=.031*	.40 S=.001*	.65 S=.001*		.21 S=.045*	.50 S=.001*	.13 S=.153	.25 S=.023*	.30 S=.007*	.27 S=.015*
P ARR	.39 S=.001*	.44 S=.001*	.59 S=.001*	.21 S=.045*		.49 S=.001*	.25 S=.022*	.37 S=.001*	.31 S=.006*	.37 S=.001*
Blk D	.42 S=.001*	.57 S=.001*	.83 S=.001*	.50 S=.001*	.49 S=.001*		.42 S=.001*	.49 S=.001*	.50 S=.001*	.52 S=.001*
CODE	.04 S=.385	.14 S=.134	.56 S=.001*	.13 S=.153	.25 S=.022*	.42 S=.001*		.25 S=.021*	.22 S=.041*	.25 S=.024*
CAT V	.77 S=.001*	.57 S=.001*	.38 S=.001*	.25 S=.023*	.37 S=.001*	.49 S=.001*	.25 S=.021*		.84 S=.001*	.95 S=.001*
CAT C	.77 S=.001*	.65 S=.001*	.42 S=.001*	.30 S=.007*	.31 S=.006*	.50 S=.001*	.22 S=.041*	.84 S=.001*		.95 S=.001*
CAT ALL	.80 S=.001*	.64 S=.001*	.42 S=.001*	.27 S=.015*	.37 S=.001*	.52 S=.001*	.25 S=.024*	.95 S=.001*	.95 S=.001*	

* Significant at .05 or better

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE XVI

Kendall Correlation Coefficients and Significance Levels on All Variables Within the Study (N=65)

VARIABLE GUIDE	VARIABLE PAIR	VARIABLE PAIR	VARIABLE PAIR	VARIABLE PAIR	VARIABLE PAIR	VARIABLE PAIR
	AGE W/GRADE t=.83 S=.001*	AGE W/NUM SC t=.00 S=.498	AGE W/SC t=.03 S-.397	AGE W/BITCH t=.23 S-.009*	AGE W/FIQ t=-.02 S=.420	AGE W/VIQ t=-.04 S=.323
A G E	AGE W/INFO t=-.07 S=.230	AGE W/SIM t=-.01 S=.478	AGE W/ARITH t=-.02 S=.443	AGE W/VOCAB t=-.10 S=.151	AGE W/COMP t=.01 S=.467	AGE W/PIQ t=.03 S=.380
	AGE W/P COMP t=.19 S=.028.	AGE W/P ARR t=-.02 S=.426	AGE W/BLK D t=.00 S=.488	AGE W/COD t=-.01 S=.455	AGE W/CAT-V t=.05 S=.299	AGE W/CAT-C t=.09 S=.190
	AGE W/CAT-A t=.07 S=.235	GR W/NUM SC t=.15 S=.063	GR W/SC t=.17 S=.069	GR W/BITCH t=.24 S=.006*	GR W/FIQ t=.13 S=.086	GRADE W/VIQ t=.10 S=.148
G R A D E	GR W/INFO t=.06 S=.257	GR W/SIM t=.10 S=.161	GR W/ARITH t=.11 S=.128	GR W/VOCAB t=.04 S=.349	GR W/COMP t=.09 S=.170	GR W/PIQ t=.14 S=.070
	GR W/P COM t=.30 S=.001*	GR W/P ARR t=.04 S=.338	GR W/BLK D t=.12 S=.113	GR W/CODE t=.04 S=.330	GR W/CAT-V t=.17 S=.043	GR W/CAT-C t=.20 S=.018*
N U M	GR W CAT-A t=.18 S=.028*	NUM SC W/SC t=.72 S=.001*	NUM SC W/BITCH t=.19 S=.016*	NUM SC W/FIQ t=.40 S=.001*	NUM SC W/VIQ t=.37 S=.001*	NUM SC W/INFO t=.40 S=.001*
S C	NUM SC W/SIM t=.32 S=.001*	NUM SC W/AR t=.27 S=.002*	NUM W/VOCAB t=.40 S=.001*	NUM SC W/COM t=.31 S=.001*	NUM SC W/PIQ t=.34 S=.001*	NUM SC W/PCMP t=.22 S=.006*

*Significant at .05 or better

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE XVI (CONTINUED)

Kendall Correlation Coefficients and Significance Levels on All Variables Within the Study (N=65)

VARIABLE GUIDE	VARIABLE PAIR	VARIABLE PAIR	VARIABLE PAIR	VARIABLE PAIR	VARIABLE PAIR	VARIABLE PAIR
N						
U	NUM SC W/	NUM SC W/	NUM SC W/	NUM SC W/	NUM SC W/	NUM SC W/
M	P ARR	BLK D	CODE	CAT VOCAB	CAT COMP	CAT ALL
	t=.20	t=.37	t=.28	t=.49	t=.46	t=.47
S	S=.014*	S=.001*	S=.001*	S-.001*	S=.001*	S=.001*
C						
	SC W/BITCH	SC W/FIQ	SC W/VIQ	SC W/INFO	SC W/SIM	SC W/ARITH
	t=.11	t=.39	t=.33	t=.34	t=.30	t=.28
	S=.135	S=.001*	S=.001*	S=.001*	S=.002*	S=.004*
	SC W/VOCAB	SC W/COMP	SC W/PIQ	SC W/P COMP	SC w/P ARR	SC W/BLK D
	t=.31	t=.27	t=.38	t=.22	t=.21	t=.43
	S=.002*	S=.005*	S=.001*	S=.021*	S=.026*	S=.001*
S						
C	SC W/CODE	SC W/CAT-V	SC W/CAT-C	SC W/CAT-ALL	BITCH W/FIQ	BITCH W/FIQ
	t=.34	t=.38	t=.42	t=.40	t=.15	t=.20
	S=.001*	S=.001*	S=.001*	S=.001*	S=.040*	S=.010*
	BITCH W/IN	BITCH W/SIM	BITCH W/ARR	BITCH W/VOC	BITCH W/COM	BITCH W/PIQ
	t=.23	t=.16	t=.18	t=/31	t=.14	t=.09
	S=.006*	S=.037*	S=.024*	S=.001*	S-.061	S-.157
B						
I	BITCH W/P C	BITCH W/P A	BITCH W/B D	BITCH W/CODE	BITCH W/CAT-V	BITCH W/CAT-C
T	t=.01	t=.16	t=.08	t=.06	t=.36	t=.28
C	S=.446	S=.039*	S=.176	S=.235	S=.001*	S=.001*
H						
	BITCH W/CAT-A	FIQ W/VIQ	FIQ W/INFO	FIQ/WSIM	FIQ W/ARITH	FIQ W/VPC
	t=.36	t=.76	t=.60	t=.66	t=.53	t=.53
	S=.001*	S=.001*	S=.001*	S=.001*	S=.001*	S=.001*
F	FIQ W/COMP	FIQ W/PIQ	FIQ W/P C	FIQ W/P ARR	FIQ W/BLK D	FIQ W/CODE
I	t=.65	t=.72	t=.45	t=.43	t=.61	t=-.28
Q	S=.001*	S=.001*	S=.001*	S=.001*	S=.001*	S-.001*

* Significant at .05 or better

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE XVI (CONTINUED)

Kendall Correlation Coefficients and Significance Levels on All Variables Within the Study (N=65)

VARIABLE GUIDE	VARIABLE PAIR	VARIABLE PAIR	VARIABLE PAIR	VARIABLE PAIR	VARIABLE PAIR	VARIABLE PAIR
F	FIQ W/CAT-V	FIQ W/CAT-C	FIQ W/CAT-A	VIQ W/INFO	VIQ W/SIM	VIQ/W ARITH
S	t=.44	t=-.49	t=.48	t=.73	t=.73	t=.54
I	S=.001*	S=.001*	S=.001*	S=.001*	S=.001*	S=.001*
Q	VIQ W/VOC	VIQ W/COMP	VIQ W/PIQ	VIQ W/P C	VIQ W/P ARR	VIQ W/BLK D
	t=.68	t=.74	t=.46	t=.34	t=.35	t=.44
V	S=.001*	S=.001*	S=.001*	S=.001*	S=.001*	S=.001*
I						
Q	VIQ W/CODE	VIQ W/CAT-V	VIQ W/CAT-C	VIQ W/CAT-A	INFO W/SIM	INFO W/ARITH
	t=.13	t=.50	t=.56	t=.55	t=.56	5=.43
	S=.075	S=.001*	S=.001*	S=.001*	S=.001*	S=.001*
	INFO W/VOC	INFO W/COMP	INFO W/PIQ	INFO W/P COM	INFO W/P ARR	INFO W/BLK D
	t=.69	t=.54	t=.35	t=.25	t=.31	t=.35
I	S=.001*	S=.001*	S=.001*	S=.003*	S=.001*	S=.001*
N						
F	INFO W/CODE	INFO W/CAT-V	INFO W/CAT-C	INFO W/CAT-A	SIM W/ARITH	SIM W/VOCAB
O	t=.15	t=.57	t=.58	t=.62	t=.39	t=.52
	S=.049*	S=.001*	S=.001*	S=.001*	S=.001*	S=.001*
	SIM W/COMP	SIM W/PIQ	SIM W/P COM	SIM W/P ARR	SIM W/BLK D	SIM W/CODE
	t=.61	t=.45	t=.31	t=.36	t=.40	t=.16
S	S=.001*	S=.001*	S=.001*	S=.001*	S=.001*	S=.040*
I						
M	SIM W/CAT-V	SIM W/CAT-C	SIM W/CAT-A	ARITH W/VOC	ARITH W/COM	ARITH W/PIQ
	t=.35	t=.44	t=.40	t=.46	t=.45	t=.39
	S=.001*	S=.001*	S=.001*	S=.001*	S=.001*	S=.001*
A						
R	ARITH W P C	ARITH W P AR	ARITH W B D	ARITH W COD	ARITH W/CAT-V	ARITH W/CAT-C
I	t=.32	t=.20	t=.40	t=.16	t=.28	t=.31
T	S=.001*	S=.017*	S=.001*	S=.041*	S=.001*	S=.001*
H						

* Significant at .05 or better

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE XVI (CONTINUED)

Kendall Correlation Coefficients and Significance Levels on All Variables Within the Study (N=65)

VARIABLE GUIDE	VARIABLE PAIR	VARIABLE PAIR	VARIABLE PAIR	VARIABLE PAIR	VARIABLE PAIR	VARIABLE PAIR
ARITH	AR W/CAT-A t=.31 S=.001*	VOC W/COMP t=.50 S=.001*	VOC W/PIQ t=.31 S=.001*	VOC W/P COM t=.22 S=.009*	VOC W/P ARR t=.32 S=.001*	VOC W/BLK D t=.33 S=.001*
V O C A B	VOC W/CODE t=.06 S=.249	VOC W/CAT-V t=.62 S=.001*	VOC W/CAT-C t=.61 S=.001*	VOC W/CAT-A t=.65 S=.001*	COMP W/PIQ t=.43 S=.001*	COMP W/P COMP t=.31 S=.001*
COMP	COMP W/P A t=.33 S=.001*	COMP W/BLK D t=.43 S=.001*	COMP W/CODE t=.10 S=.134	COMP W/CAT-V t=.45 S=.001*	COMP W/CAT-C t=.48 S=.001*	COMP W/CAT-A t=.48 S=.001*
P I Q	PIQ W/P COM t=.49 S=.001*	PIQ W/P ARR t=.45 S=.001*	PIQ W/BLK D t=.71 S-.001*	PIQ W/CODE t=.44 S=.001*	PIQ W/CAT-V t=.27 S=.001*	PIQ W/CAT-C t=.32 S=.001*
P C	PIQ W/CAT-A t=.30 S=.001*	P C W/P ARR t=.14 S=.069	P C W/BLK D t=.39 S=.001*	P C W/CODE t=.11 S=.107	P C W/CAT-V t=.18 S-.024*	P C W/CAT-C t=.25 S=.002*
P A	P C W/CAT-A t=.22 S=.008*	P A W/BLK D t=.40 S=.001*	P A W/CODE t=.26 S=.002*	P A W/CAT-V t=.28 S=.001*	P A W/CAT-C t=.26 S=.002*	P A W/CAT-A t=.28 S=.001*
B C D & O D	B D W/CODE t=.36 S=.001*	B D W/CAT-V t=.31 S=.001*	B D W/CAT-C t=.36 S=.001*	B D W/CAT-A t=.33 S=.001*	CODE W/CAT-V t=.15 S=.051*	CODE W/CAT-C t=.20 S=.018*
C A T	CODE W CAT-A t=.18 S=.024*	CAT-V W/CAT-C t=.66 S=.001*	CAT-C W/CAT-A t=.85 S=.001*	CAT-C W/CAT-A t=.82 S=.001*		

* Significant at .05 or better

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE XVII

Analysis of Variance Scores, F Statistic, and Significance Levels Between Social Class (By Category) and Study Variables (N=65)

CRITERION VARIABLE	SOURCE	SQUARES SUMS	D.F.	MEAN SQUARE	F	SIGNIFICANCE
NUMERICAL(WEIGHTED)SC	Between groups	4246.698	1	4246.698	230.18	.0001*
	Within groups	1162.318	63	18.449		
BITCH	Between groups	121.237	1	121.237	0.90	.3459
	Within groups	8469.009	63	134.429		
FULL I.Q.	Between groups	3900.994	1	3900.994	19.441	.0001*
	Within groups	12641.560	63	200.00		
VERBAL I.Q.	Between groups	2949.441	1	2949.441	12.20	.0009*
	Within groups	15228.713	63	241.726		
INFORMATION (WISC-R)	Between groups	104.134	1	104.134	11.24	.0014*
	Within groups	583.620	63	9.264		
SIMILARITIES (WISC-R)	Between groups	111.029	1	111.029	9.11	.0037*
	Within groups	767.833	63	12.188		
ARITHMETIC (WISC-R)	Between groups	54.483	1	54.483	9.14	.0036*
	Within groups	375.302	63	5.957		
VOCABULARY (WISC-R)	Between groups	85.907	1	85.907	10.23	.0022*
	Within groups	529.078	63	8.398		
COMPREHENSION (WISC-R)	Between groups	95.211	1	95.211	7.069	.0099*
	Within groups	848.543	63	13.469		
PERFORMANCE I.Q. (WISC-R)	Between groups	3429.839	1	3429.839	17.647	.0001*
	Within groups	12244.223	63	194.353		
PICTURE COMPLETION (WISC-R)	Between groups	58.094	1	58.094	5.140	.0268*
	Within groups	712.060	63	11.303		
PICTURE ARRANGEMENT (WISC-R)	Between groups	34.707	1	34.707	4.011	.0495*
	Within groups	545.078	63	8.652		
BLOCK DESIGN (WISC-R)	Between groups	115.419	1	115.419	19.715	.0001*
	Within groups	368.827	63	5.854		

* Significant at .05 or better

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE XVII (CONTINUED)

Analysis of Variance Scores, F Statistic, and Significance Levels Between Social Class (By Category) and Study Variables (N=65)

CRITERION VARIABLE	SOURCE	SQUARES SUMS	D.F.	MEAN SQUARE	F	SIGNIFICANCE
CODING (WISC-R)	Between groups	117.646	1	117.646	9.89	.0025*
	Within groups	749.339	63	11.894		
CALIF. ACHIEVEMENT TEST (VOCABULARY)	Between groups	144.419	1	144.419	16.65	.0001*
	Wtihin groups	546.394	63	8.673		
CALIF. ACHIEVEMENT TEST (COMPREHENSION)	Between groups	162.383	1	162.383	19.60	.0001*
	Within groups	522.051	63	8.287		
CALIF. ACHIEVEMENT TEST (TOTAL READING SCORE)	Between groups	151.812	1	151.812	18.99	.0000*
	Within groups	503.784	63	7.997		

* Significant at .05 or better

TABLE XVIII

T-Test Results to Determine Significance Between Male/Female Sample Means

VARIABLE	# OF CASES	MEAN	S.D.	STD.ERROR	F VALUE	D.F.	TWO-TAIL PROB.
BITCH MALE	31	62.84	10.49	1.88	1.44	63	.315
BITCH FEMALE	34	61.03	12.60	2.16			

CHAPTER V

DISCUSSION OF RESULTS

Introduction

This chapter will address itself to a discussion of the relevance of the data as it bears upon the hypotheses under consideration in this study. In particular, statistical evidence will be produced which allows for either rejection or confirmation of H_0I, H_0II and H_0III. Additional findings and their implications will be considered, along with possible followup suggested by the data base produced by the participants in this research. Limitations of the study will be addressed, followed by a concluding summary of the findings.

Data Relevance and Its Relation To the Study's Hypotheses

Hypothesis I: In its null form, this hypothesis indicates that there is no relationship between the BITCH and the Verbal scale of the WISC-R. The statistical evidence strongly refutes this statement. Pearson correlation coefficients between the BITCH and WISC-R Verbal I.Q. measures ($r=.26$) for this sample ($N=65$; $df=63$) indicate a significance level of .017 ($.01 < P < .05$). The null hypothesis can therefore be rejected at the .05 level.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Thus the alternate hypothesis supporting a significant relationship between these variables can be confirmed. It is interesting to note that all WISC-R Verbal I.Q. subtests achieve significance at the .003 confidence level ($r=.34$; $P \prec .01$). It appears reasonable to conclude that since the vocabulary subtest has the highest correlation with the BITCH and achieves significance at well beyond the .01 level, a common factor relating to both instruments may well be a comprehension of vocabulary. This would be in line with factor analytic studies which indicate that of three factors (verbal comprehension, perceptual organization, freedom of distractability) which load the WISC-R, the verbal comprehension factor most highly saturates the Verbal scale of the instrument.[1] Certainly whatever the Verbal portion of the Wechsler (especially the vocabulary subtest) measures, is also measured in part by the BITCH, as supported by the statistical results. With reference to the WISC-R Performance (non-verbal) scale, correlations with the BITCH were generally non-significant, with the exception of the Picture Arrangement subtest which yielded a significance level of .047 ($r=.21$ $P \prec .05$). In examining this particular subtest, one finds that the picture arrangement sequencing required by the test taker is preceded by much verbal explanation and rationale. Although at this point only conjecture is possible as to why this particular non-verbal test should achieve a higher level of significance, it seems to this writer that the verbal comprehension required of an individual

---

[1] Kaufman, Op. cit.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

to adequately handle this subtest also renders it more likely to correlate with the BITCH. Interestingly, the Picture Arrangement subtest correlates highly with the WISC-R Verbal I.Q. ($r=.48$; $p < .001$).

Hypothesis II: H_oII posits no relationship between the BITCH and verbal achievement. In this instances, the measure of verbal achievement was the Reading section of the CAT comprising two subtests (Vocabulary and Comprehension). Correlations between the BITCH and the CAT total reading score attain significance at the .001 level ($r=.44$; $p < .001$); this is also the case for the CAT Vocabulary subtest ($r=.47$; $p < .001$). The CAT Reading Comprehension portion was almost as great ($r=.35$; $p < .002$). The null hypothesis can be rejected with confidence and the alternative hypothesis bearing on the significant relationship between the BITCH and verbal achievement (as measured by the CAT) can be confirmed.

Hypothesis III: H_oIII states that there is no relationship in scores on the BITCH between middle and lower class samples. In contrast, not only does H_1III specify a significant relationship between Black social class groupings, but also posits directionality, with higher scores being made by lower SES Blacks. It must be pointed out that two measures of SES were employed in this study, specifically a weighted social class measure (numerical SC) and social class by category (i.e., lower SES "category one" and middle SES "category two" groups). In comparing scores of the two groups by SES category no significant relationship was seen on either of two

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

correlational measures Pearson "r" and Kendall "t" ($r=.12$; $P > .05$; $t=.11$; $P > .05$). However, when the weighted SES scores were compared on this variable, they achieved significance ($r=.23$; $.01 < P < .05$; $t=.19$; $.01 < P < .05$). Thus in terms of weighted social class scores, H_oIII can be rejected at the .05 level. Although rejection of the null hypothesis is accomplished at the .05 level, in this instance H_1III cannot be confirmed. In comparing the BITCH mean scores between the two SES groups, it is seen that the middle social class sample scored somewhat higher (M=63.3) than their lower SES counterparts (M=60.58). Accordingly, if a significant correlation in favor of the lower were to be achieved, it would have to be a negative one; this is not supported by the data. In other words, although there is a significant relationship at greater than the .05 level of confidence, it is in the opposite direction; middle social class youngsters in this sample scored significantly higher on the BITCH than did the lower social class sample. Thus the alternate hypothesis contending that lower SES groups will score higher on the BITCH cannot be confirmed on the basis of the data obtained in this study. Upon closer examination, this apparent contradiction is not too surprising.

The BITCH correlates significantly with both the WISC-R Verbal Scales and the California Achievement Test (Reading). Frequently quoted research findings support the fact that higher socioeconomic class is positively correlated with I.Q. and achievement measures (Kaufman, 1973; Kaufman and Doppelt, 1976). Therefore middle class youngsters obtaining higher BITCH scores

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

than the lower class group are entirely consistent with other studies relating social class to intelligence and achievement measures.

Additional Findings

At this point, it would be well to review other findings relevant to the BITCH that are supported by the data obtained in this study. Comparison of the BITCH with age indices show a significant relationship ($r=.29$; $P < .01$). Correlational measures between the BITCH and grade variables are also interesting ($r=.34$; $P < .01$). Thus within the parameters of this study, the youngsters who were older and in higher grades did better than the others on the BITCH. A comparison of BITCH scores by sex did not yield significant results; a T-test computed on sample means regarding this variable did not support the possibility of either males or females obtaining better scores (T value=0.63; $P > .05$). Although the BITCH correlated significantly with the WISC-R on Full I.Q. ($r=.22$; $.01 < P < .05$), the level of significance was appreciably greater for Verbal I.Q. ($P < .017$). In spite of this finding, the Comprehension subscale on the verbal portion of the WISC-R yielded an insignificant relationship with the BITCH ($r=.17$; $P > .05$). Since the Comprehension subtest "taps" verbally mediated "common sense" reasoning items, the possibility that the BITCH may be a rather narrow measure of verbal ability must be considered.

At this juncture, a comparison of BITCH means between this sample (61.9) and Williams' original Black group in St.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Louis (87.1) is in order. A $\overline{Z}$ test (Critical Ratio for large samples) was computed. With a combined N of 165, this statistic seemed particularly appropriate for determination of significant differences between the sample groups. The results ($\overline{Z}$=15.75; $P < .001$) clearly demonstrate the significance of these differences in mean BITCH scores between the St. Louis sample and the Ss in the present study.

Another finding which must be mentioned deals with the WISC-R scores derived from the participants in the study. The mean I. Q. score for both groups was 102.33 with an S.D. of 16.1; this examiner did not find the discrepancy of about one S.D. which has been reported in other research. An inspection of WISC-R scores by social class does show a difference of about one standard deviation (fifteen points) which proved to be quite significant. An Analysis of Variance computed on this data yielded an F Score of 19.44 ($P < .0001$). Indeed on all major test variables, including scores on the BITCH, WISC-R, and CAT (Reading) weighted social class proved to be a significant discriminating factor relevant to how well one did in answering test-related items.

Up to this time, little has been stated about the WISC-R results within the sample. Not only did it correlate very significantly with social class (Categorical SES: r=.49; $P < .001$; Weighted SES: r=.57; $P < .001$), but also with all achievement measures. This is best indicated by the relationship between the WISC-R and the CAT (total reading) instrument (r=66; $P < .001$). A comparison of the differences

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

between the WISC-R and the BITCH in relation to the CAT, can be seen in the fact that a Coefficient of Determination (r^2) would indicate that while the BITCH only accounts for about 20% of the variance held in common by these instruments, the WISC-R shares about 42% variance with the CAT. Thus the WISC-R would seem to be about twice as powerful a predictor of achievement criteria than would the BITCH.

Implications and Conclusions

In light of the findings supported by the results obtained from the study, what implications can be drawn and what conclusions can be reached? It appears to this examiner that the foregoing statements form a close connection with the data to wit:

I. The BITCH is significantly correlated with the WISC-R, a well-established measure of intelligence that includes a normative Black sample in proportion to the prevailing census figures.

II. The BITCH is significantly correlated with the CAT (Reading) and on the surface appears to have some predictive power in relation to achievement measures (as indicated by the CAT). However, the WISC-R's correlational base is considerably higher than that of the BITCH, so that it is a measure with far greater predictive power for achievement criteria than the BITCH.

III. Although the BITCH does in fact very modestly discriminate between social class groups, it apparently does so in favor of the higher SES group; this may reflect negatively

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

upon some of its theoretical assumptions.

IV. The BITCH significantly correlates with both age and grade; its use for younger groups or for the earlier grades is suspect.

V. A comparison of BITCH means between participants in this study and the original norming group in St. Louis, Mo. shows very significant differences in favor of the original group. The following possibilities must therefore be considered:

1. A marked regional difference exists between these groups.
2. A considerable change in "slang" usage has occured with time. (Note: Williams' original test materials and data are now six years old.)

VI. A comparison of the BITCH with all WISC-R Verbal subtests show a moderately significant relationship. Results obtained from the Comprehension subtest on the WISC-R were, however, insignificant. This suggests that the BITCH may well be a more narrowly-based measure of ability than is the case with the WISC-R.

VII. The generalized finding that Blacks are about one Standard Deviation below the population mean for the United States was not borne out in this study. There was, however, a social class difference of about one Standard Deviation in favor of the middle-class group. Perhaps the selection of a "typical" (i.e., school-attending, grade-appropriate) sample has much to do with measured mental ability.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

VIII. Since a major complaint of the culture-specific test movement relates to the inappropriate (i.e., biased) usage of norm-referenced instruments (e.g., the WISC-R) for racial or ethnic minorities, this study tends to refute that contention. Indeed, it demonstrates the predictive power of such an instrument when appropriate norms are included in the standardization sample.

IX. Finally, it appears that a rather narrowly-based culture-specific test, such as the BITCH, doesn't seem to add much predictive power, if any, to the instruments currently in use.

Limitations of the Study

Although this study was carefully designed and executed, it is but a single research project with only a moderately-sized sample. It is obvious that considerable replication of its findings remains to be accomplished if its conclusions are to be generalized to larger populations. Moreover, the study was accomplished in a region far removed from the original locale of the BITCH. Obviously, it is very necessary to mount similar research in other regions, so as to shed light upon the most appropriate instruments to be used in screening minority youngsters in terms of mental ability or academic achievement.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Concluding Summary

This paper has defined a problem which grew out of a felt need by several spokesmen to refute the notion that Black populations were natively inferior to other populations in terms of mental ability. Although a considerable body of research has been initiated in this area during the past five decades, more recent investigators have sparked bitter debate; their findings have been subject to merciless scrutiny and criticism, especially in light of the tendency of the media to popularize or pull out of context controversial issues. In recent years, few such issues have been so heatedly argued as has been this one. The culture-specific test movement emerged from the sincere desire of some researchers to show that biased tests unfairly discriminate segments of the populations. This investigator certainly agrees, but, having demonstrated this point, are the claims made by the BITCH and similar instruments valid and able to be generalized, or not? It is in the interest of the individual, as well as science, to determine this truth on the basis of the evidence, and not on the basis of emotion, no matter how well-intended. This project was initiated for this purpose; the author hopes it has succeeded in this endeavor.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

BIBLIOGRAPHY

Books

Anastasi, Anne. Psychological Testing, 4th Ed., New York: Mac Millan, 1976.

Anastasi, Anne. Testing Problems in Perspective, Wash. D.C.: American Council on Education, 1965.

Baxter, Paul. Race and Social Difference, Harmondworth: Penguin, 1972.

Beck, S.J. and Molish, H. (Eds.), Reflexes to Intelligence, Glencoe, Ill.: Free Press, 1959.

Blake, R.R. and Ramsey, G.V. Perception: An Approach to Personality, New York: Ronald, 1951

Block, N.J. and Dworkin, Gerald. The I.Q. Controversy: Critical Readings, New York: Pantheon, 1976.

Bowles, Samuel. Schooling and Inequality from Generation to Generation, Cambridge, Mass.: Center for Educational Policy Research, 1971.

Brace, C.L., Gamble, G.R., Bond, J.T. (Eds.). Race and Intelligence, Wash, D.C.: American Anthropology Assn., 1971.

Brigham, C.C. A Study of American Intelligence, Princeton, N.J., Princeton University Press, 1923.

Brim, Orville, et al. Experiences and Attitudes of American Adults Concerning Standardized Intelligence Tests, Tech. Rep. #1, Social Consequences of Testing, New York: Russell Sage Foundation, 1965.

Buros, O.K. (Ed.). Seventh Mental Measurements Yearbook, Highland Park, N.J.: Gryphon, 1972.

Butcher, H.J. Human Intelligence: Its Nature and Assessment, London: 1968.

Callahan, Raymond. Education and the Cult of Efficiency, Chicago: University of Chicago Press, 1962.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Campbell, D.T., Stanley, J.C. Experimental and Quasi-Experimental Designs for Research, Chicago: Rand-McNally, 1963.

Cancro, R. (Ed). Intelligence: Genetic and Environmental Influences, New York: Grune and Stratton, 1971.

Cattell, R.B. Abilities: Their Structure, Growth and Action, Boston: Houghton-Mifflin, 1971.

Cronbach, Lee J. Essentials of Psychological Testing, New York: Harper, 1949.

Davis, A., Eells, K. Davis-Eells Test of General Intelligence or Problem-Solving Ability Manual, New York: World Book Co., 1953.

Davis, Allison. Social Class Influences on Learning, Cambridge, Mass.: Harvard University Press, 1948.

Dobzhansky, Th. The Genetics of the Evolutionary Process, New York: Columbia University Press, 1970.

Dockrell, W.D. (Ed.). Toronto Symposium on Intelligence, Toronto: Ontario Institute for Studies in Education, 1970.

Duberman, Lucile. Social Inequality, Philadelphia: Lippincott, 1976.

Eells, Kenneth, Davis, Allison, et al. Intelligence and Cultural Difference, Chicago: University of Chicago Press, 1951.

Ehrlich, Paul and Feldman, Shipley. The Race Bomb, New York: New York Times Book Co., 1977.

Eysenck, H.J. Inequality of Man, London: Temple-Smith, 1973.

Eysenck, H.J. Measurement of Intelligence, Baltimore: Williams and Wilkins, 1973.

Eysenck, Hans. Race, Intelligence and Education, London: Temple-Smith, 1971.

Fincher, Jack. Human Intelligence, New York: Putnam, 1976.

Forbes, J. Afro-Americans in the Far West, Berkeley: Far West Labs for Educational Research and Development, 1967.

Ginsburh, H. Myth of the Deprived Child, Englewood Cliffs, N.J.: Prentice-Hall, 1972.

Gronlund, N.E. Measurement and Evaluation in Teaching, 2nd Ed., New York: Mac Millan, 1971.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Hardyck, C.D. and Petrinovich, L.F. Statistics for the Behavioral Sciences, Philadelphia: Saunders, 1976.

Herrnstein, R.J. I.Q. in the Meritocracy, Boston: Little-Brown, 1971.

Hofstadter, Richard. Social Darwinism in American Thought, Rev. Ed., New York: Braziller, 1969.

Hollingshead, August and Redlich, Frederick. Social Class and Mental Illness, New York: John Wiley & Sons, 1958.

Hove, Nathan. The Black Anglo-Saxons, New York: Marzini & Munsell, 1965.

Issac, S. and Michael, W. Handbook in Research and Evaluation, San Diego: Edits Publishes, 1971.

Jensen, A.R. Educability and Group Differences, New York: Harper & Row, 1973.

Jensen, Arthur. Genetics and Education, New York: Harper & Row, 1972.

Johnson, Orval. Tests and Measurement in Child Development: A Handbook, San Francisco: Jossey-Bass, 1976.

Kamin, L.J. The Science and Politics of I.Q., New York: Wiley, 1974.

Kennedy, Wallace. A Followup Study of Negro Intelligence and Achievement, Chicago: University of Chicago Press, 1969.

Kirk, Roger, Experimental Design: Procedures for the Behavioral Sciences, Belmont, Ca.: Brooks-Cole, 1968.

Kleka, W.R., Nie, H.H., Hull, C.D. S.P.S.S. Primer, New York: McGraw-Hill, 1975.

Klineberg, Otto (Ed.). Characteristics of the American Negro, New York: Harper & Row, 1944.

Kohn, Melvin. Class and Conformity, Ill.: Dorsey Press, 1969.

Labov, W. Language in the Inner City: Studies in Black English Vernacular, Philadelphia: University Pennsylvania Press, 1972.

Loehlin, John. Race Difference in Intelligence, San Francisco: W.H. Freeman, 1975.

Major, C. Dictionary of Afro-American Slang, New York: International Publishers, 1970.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Maranell, Gary. Scaling: A Sourcebook for the Behavioral Sciences, Chicago: Aldine, 1974.

Mercer, Jane and Lewis, June. SOMPA Technical Manual, New York: Psychological Corp., 1977.

Nie, N.H. et al. S.P.S.S.: Statistical Package for the Social Science Manual, 2nd Ed., New York: McGraw-Hill, 1975.

Oakland, Thomas, Psychological and Educational Assessment of Minority Children, New York: Bruner-Mazel, 1977.

Pettigrew, Thomas. A Profile of the Negro American, Princeton, N.J.: Van Nostrand, 1964.

_____, Publication Manual of the American Psychological Assn., (2nd Ed.), Baltimore: Pridework Press, 1974.

Richardson, K. and Spears, D. (Eds.). Race and Intelligence, Baltimore: Penguin, 1972.

Richardson, K. and Spears, D. (Eds.). Race, Culture and Intelligence, Baltimore: Penguin, 1972.

Reiss, A.J. Occupations and Social Status, New York: Free Press, 1961.

Salvia, J. and Ysseldyke, J. Assessment in Special and Remedial Education, Boston: Houghton-Mifflin, 1978.

Samuda, Ronald. Psychological Testing of American Minorities, New York: Harper & Row, 1975.

Sattler, J. Assessment of Children's Intelligence, Philadelphia: Saunders, 1974.

Senna, Carl (Ed.). The Fallacy of I.Q., New York: Third Press 1973.

Shuey, Audrey. The Testing of Negro Intelligence, 2nd Ed., New York: Social Science Press, 1966.

Siegel, Sidney. Nonparametric Statistics, New York: McGraw-Hill, 1956.

_____. Supplement to the Dictionary of Occupational Titles, 3rd Ed., Wash, D.C.: U.S. Government Printing Office, 1966.

Tiegs, Ernest & Clark, Willis W., California Achievement Tests, 1970 Ed., Level 5 Form A, Manual, Monterey, Ca.: Mc Graw-Hill/CTB, 1970.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Terman, Lewis, Merrill, Maud. Measuring Intelligence, Boston: Houghton-Mifflin, 1937.

Tocqueville, Alexis. Democracy in America, (Trans. & Revised), New York: Oxford University Press, 1947.

Turabian, K.L. A Manual for Writers of Term Papers, Theses, and Dissertations, 4th Ed., Chicago, University of Chicago Press, 1973.

U.S. Bureau of the Census, Census of Population: 1970 General Population Characteristics, Final Report PC(1)B1, United States Summary, Wash, D.C.: U.S. Govt. Printing Office, 1972.

Warner, W.L. Social Class in America: A Manual of Procedure for Measuring Social Class, Gloucester, Mass.: P. Smith, 1957.

Wechsler, David. The Measurement of Adult Intelligence, Baltimore: Williams & Wilkins, 1958.

Wechsler, David. WISC Manual, New York: Psychological Corp., 1949.

Wechsler, David. WISC-R Manual, New York: Psychological Corp., 1974.

Whimbey, Arthur. Intelligence Can Be Taught, New York: Dutton, 1975.

Williams, R.L. Black Intelligence Test of Cultural Homogeneity, Test Manual, St. Louis: Williams & Assoc., 1972.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Articles

Alfert, E. "Comment on the Promotion of Prejudice." Journal of Social Issues, V. 25(4), (1969), pp. 206-211.

Baldwin, T.L., McFarlane, P.T., Garvey, C.L. "Children's Communication Accuracy Related to Race and Socioeconomic Status." Child Development, V. 42, (1971), pp. 345-357.

Barnesbey, N.S. "The Effect of the Race of the Examiner on the Test Performance of Negro and White Children," Proc. of 81st Ann. Conv. of American Psychological Assn., Montreal, Canada, V. 8, (1973), pp. 647-648.

Bernal, E.M. "A Response to Educational Uses of Tests and Disadvantaged Students." American Psychologist, V. 30 (1975), pp. 93-95.

Binet, A. and Henri, V. "La Psychologic Individuelle." Annee Psychologique, V. 2, (1895), pp. 411-463. Quoted in Lewis Terman, "The Mental Test as a Psychological Method," Psychology Review, V. 31, (1924), pp. 93-117.

Boehm, Ann. "Criteria-Referenced Assessment for the Teacher." Teacher College Record, V. 75, (1973), pp. 117-126.

Boehm, V.R. "Negro-White Differences in Validity of Employment and Training Selection Procedures: Summary of Recent Evidence." Journal Apl. Psych., V. 56, (1972), pp. 33-39.

Boone, J.A. and Adesso, V.J. "Racial Differences on a Black Intelligence Test." Journal of Negro Education, Fall 1974, pp. 429-435.

Breiter, C. "The Relation of Social Class and Educational Aptitude." Canad. Psychol., V. 13, (1972), pp. 329-340.

Brigham, C.C. "Value of Tests in Examinations of Immigrants." Industrial Psych., V. 1, (1926), pp. 412-417.

Buck, Mildred. "The Multi-Dimensional Model for the Assessment of Children Referred for Classes for Mental Retardation." Journal of Afro-American Issues, V. 3 #1, (1975), pp. 91-101.

Burt, Cyril. "The Genetic Determination of Differences in Intelligence: A Study of MZ Twins Reared Together and Apart." British Journal of Psychology, V. 57, (1966), pp. 137-153.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Burnes, E. "I. Q. Testing and Minority Children: Imperatives for Change." National Leadership Inst. Teacher Educ., University of Connecticut Tech. Paper, pp. 1-8.

Burns, Kay. "Patterns of WISC Scores for Children of Two Socioeconomic Classes and Races." Child Development, V. 41, (1970), pp. 493-499.

Caldwell, M.B. & Knight, D. "The Effect of Negro and White Examiners on Negro Intelligence Test Performance," Journal of Negro Education, V. 39, (1970), pp. 177-179.

Campbell, J.T. "Tests are Valid for Minority Groups Too." Publ. Personnel Management, V. 2 (1973), pp. 70-73.

Cattell, R.D. "A Culture-Free Intelligence Test." Journal Of Educ. Psych., V. 31, (1940), pp. 161-179.

Cleary, T.A., et al. "Educational Uses of Tests with Disadvantaged Students." American Psychologist, V. 30, (1975), pp. 15-41.

Cronbach, Lee. "Five Decades of Public Controversy over Mental Testing." American Psychologist, V. 30, (1975), pp. 1-14.

Cronbach, Lee J. "Beyond the Two Disciplines of Scientific Psychology." American Psychologist, V. 30, (1975), pp. 116-117.

Duncan & Duncan. "A Methodological Analysis of Segregation Indexes." American Soc. Review, V. 20, (1955), pp. 210-217.

Dyer, H.S. "Is Testing a Menace to Education?" N.Y. State Educ. XLIX (Oct. 1961), pp. 16-19. In Readings in Educational and Psychological Measurement, C.I. Chase & H. G. Ludlow (Eds.), Boston: Houghton-Mifflin, (1966), pp. 40-45.

Ebel, R.L. "The Social Consequences of Educational Testing." Proceedings of the 1963 Invitational Conference on Testing Problems, Princeton, N.J.: Educational Testin Service, (1963), pp. 130-143.

Fishman, J. et al. "Guidelines for Testing Minority Group Children." Journal of Social Issues, Supplement, V. 20, (1964), pp. 129-145.

Gallagher, J.J. and Moss, J.W. "New Concepts of Intelligence and Their Effects on Exceptional Children." Exceptional Children, V. 30, (1963), pp. 1-5.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Gigliotti, R.J. "Residential Stability and Academic Sense of Control." Journal of Black Studies, V. 6, (1976), pp. 257-276.

Golden, M. et al. "Social Class Differentiation in Cognitive Development Among Balck Preschool Children." Child Development, V. 42, (1971), pp. 37-45.

Green, R.L. & Griffore, R.J., & Simmons, D. "A Restatement of the I.Q./Culture Issue." Phi Delta Kappan, (June 1976), pp. 674-676.

Grimmit, S. "The Similarities in Performance Patterns of Black and White Children: Jensen's Racial Difference Theory Questioned." Journal of Afro-American Issues, V. 4 #2, (1976), pp. 109-120.

Hall, W. S. Reder, S. & Cole, M. "Story Recall in Young Black and White Children: Effects of Racial Group Membership, Race of Examiner and Dialect." Developmental Psychology, V. 11, (1975), pp. 628-634.

Hallowell, A.I. "'Popular' Responses and Cultural Differences: An Analysis Based on Frequencies in a Group of American Indian Subjects." Rorschach Research Exchange, V. 9, (1945), pp. 153-168.

Halpen, Florence. "Clinicians Must Listen." Clin. Child Psych. Newsletter, V. 9, #4, (1970).

Harris, T.G. "Jensen, Genetics and Equalitarian Diversity." Psychology Today, (Dec. 1973), p. 40.

Hartlage, L.C. & Steele, C.T. "WISC & WISC-R Correlates of Academic Achievement." Psychology in the Schools, V. 14, (1977), pp. 15-18.

Henry, Jules. "Rorschach Technique in Primitive Cultures." American Journal of Orthopsychiatry, V. 11, (1941), pp. 230-234.

Hollingshead, A.B. & Redlich, F.C. "Social Stratification and Psychiatric Disorder." American Sociological Review, V. 18, (1953), pp. 163-169.

Horn, J.L. "Controversy Over the Nature-Nurture of Intelligence." In "Human Abilities," Annual Review of Psychology, V. 27, (1976), pp. 473-485.

Human Rights Commission. I.Q. Testing in San Francisco." Intergrateducation, Issue 63, (May-June 1973).

Humphreys, L.G., Dachler, H.P. "Jensen's Theory of Intelligence." Jour. Educational Psych., V. 60, (1969), pp. 419-433.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Humphreys, L.G. "Statistical Definitions of Test Validity for Minority Groups." Journ. App. Psych., V. 58, (1973), pp. 1-4.

Ingle, D.J. "Editorial: The Need to Study Biological Differences Among Racial Groups--Moral Issues." Perspectives in Biology and Medicine, V. 10, (1967), pp. 497-499.

Jackson, C.D. "On the Report of the Ad Hoc Committee on Education Uses of Tests with Disadvantaged Students." American Psychologist, V. 30, (1975), pp. 88-92.

Jackson, R. "Developing Criterion-Referenced Tests." ERIC Clearing-House on Tests, Measurement and Evaluation, Princeton, N.J.: Educational Testing Service, 1971.

Jacobs, J.F. & De Graaf, C.A. "Race of Subject and Race of Examiner." Integrated Education, V. 11, #4, (1973), pp. 68-71.

Jenkin, N., Spivack, G. Levine, M., & Savage, W. "Wechsler Profiles and Academic Achievement in Emotionally Disturbed Boys." Journal of Consulting Psychology, V. 28, (1964), P. 290.

Jensen, A.R. "How Much Can We Boost I.Q. and Scholastic Achievement?" Harvard Educational Review, V. 39, (1969), pp. 1-123.

Jensen, A.R. "Interaction of Level I and Level II Abilities With Race and SES." Journal Educ. Psych., V. 66, (1974), pp. 99-111.

Jensen, A.R. "I.Q. Tests Are Not Culturally Biased for Blacks and Whites." Phi Delta Kappan, (June 1976), p. 676.

Jensen, A.R. "Level I and Level II Abilities in Three Ethnic Groups." Am. Educ. Res. Jour., V. 10, (1973), pp. 263-276.

Jensen, A.R. "The Differences Are Real." Psychology Today, (Dec. 1973), pp. 79-86.

Jensen, A.R. & Figueroa, R.A. "Forward and Backward Digit Span Interaction with Race and I.Q.: Predictions From Jensen's Theory." Jour. Educ. Psych., V. 67, (1975), pp. 882-893.

Kahl, Joseph & Davis, James. "A Comparison of Indexes of Socioeconomic Status." American Sociological Review, V. 20, (1955), pp. 317-325.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Kamin, Leon. "The Misuse of I.Q. Testing." Change, (Oct. 1973), pp. 40-43.

Kaufman, A.S. "Factor Analysis of the WISC-R at 11 Age Levels Between 6 1/2 and 16 1/2 Years." Jour. of Consult. and Clin. Psych., V. 43, (1975), pp. 135-147.

Kaufman, A. S. & Doppelt, J.E. "Analysis of WISC-R Standardization Data in Terms of the Stratification Variables." Child Development, V. 47, (1976), pp. 165-171.

Lambert, Nadine & Meyers, C.E. "As a Result of a Court-Imposed Halt to I.Q. Testing in E.M.R. Placement, Children Are Being Deprived of Thier Rights to Services." Ann. Amer. Psych. Assn., V. 8, #4, (April 1977).

Lesser, G. S., Fifer, G. & Clark, D.H. "Mental Abilities of Children from Different Social Class and Cultural Groups." Mono. of Soc. for Research in Child Development, V. 30, #4, (1965).

Littell, W. M. "The Wechsler Intelligence Scale for Children: Review of a Decade of Research." Psych. Bulletin, V. 57, pp. 132-156.

Long, Peggie & Anthony, John J. "The Measurement of Mental Retardation by a Culture-Specific Test." Psychology in the Schools, (1974), pp. 310-312.

Mach Erbe, Brigitte. "Race and Socioeconomic Segregation!" American Sociological Review, V. 40 (1975, pp. 801-812.

MacQueen, R. "Larry: Case History of a Mistake." Saturday Review, (9-12-70).

McNiel, N.D. "I.Q. Tests and the Black Culture." Phi Delta Kappan, (Nov. 1975), pp. 209-210.

Meier, Dorothy & Bell, Wendell. "Anomia and Differential Access to the Achievement of Life Goals." American Sociological Review, V. 24, (1959), pp. 189-202.

Messick, S. & Anderson, S. "Educational Testing: Individual Development and Social Responsibility." Counselling Psychologist, V. 2, (1970), pp. 80-88.

Milgram, Norman. "Danger: Chauvinism, Scapegoatism and Euphemism." Clinical Child Psychology Newsletter, V. 9, #3, (Fall 1970), p. 2.

Moynihan, D.P. "Jensen Not 'Must Reading' in the Nixon Cabinet." Jour. of Social Issues, V. 26, (1970), pp. 191-192.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Nelson, C.M. & Hudson, F.G., "Prediciting the Reading Achievement of Jr. High School E.M.R. Children." American Jour. Ment. Def., V. 74, (1969), pp. 415-420.

Newland, T.E. "Testing Minority Group Children." Clin Child Psychology Newsletter, V. 9, #3, (1970).

Nuluen, F.G., Hoffman, L.J., & Bierbryer, B. "The Effects of Subject's Age, Sex, Race and SES on Psychologists' Estimates of 'True I.Q.' From WISC Scores." Jour. of Clin. Psych., V. 25 (1969), pp. 271-274.

Oakland, Thomas. "Assessing Minority Group Children: Challenges for School Psychologists." Journal of School Psychology, V. 11, #4, (1973), pp. 294-303.

Ortiz, C.C., & Ball, G. "The Enchilada Test." Inst. for Personal Effectiveness in Children, 1972.

Rivers, W. L. et al. "I. Q. Labels and Liability: Effects on the Black Child." Journal of Afro-American Issues, V. 3, #1, pp. 63-76.

Rudman, Herbert. "The Standardized Test Flap." Phil Delta Kappan, (Nov. 1977), pp. 179-185.

Samuda, R.J. "From Ethnocentrism to a Multicultural Perspective in Educational Testing." Journal of Afro-American Issues, V. 3, #1, (1975), pp.4-18.

Sapir, E. "Conceptual Categories in Primitive Languages." Science, V. 74, (1931), p. 578.

Schmidt, F.L., Herne, J.G., & Hunter, J.E. "Racial Differences in Validity of Employment Tests." Jour. Appl. Psych., V. 58, (1973), pp. 5-9.

Scott, R. Hartson, J. & Cunningham, M. "Race of Exmainer as a Variable in Test Attainments of Preschool Children." Perceptual and Motor Skills, V. 42, (1976), pp. 1167-1173.

Scriven, Michael. "The Values of the Academy: Moral Issues for American Education and Educational Research Arising from the Jensen Case." Rev. of Educ. Research, V. 40, (1970), pp. 541-549.

Silverstein, A.B. "Factor Structure of the Wechsler Intelligence Scale for Children for Three Ethnic Groups." Journ. Educ. Psych., V. 63, (1973), pp. 408-410.

Stennett, R.G. "The Relationships of Sex and Socioeconomic Class to I.Q. Change." Psychology in the Schools, V. 6, (1969), pp. 385-390.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Sowell, Thomas, "The Great I.Q. Controversy." Change, (May 1973), pp. 33-37.

Taylor, H.F. "I.Q. Heritability: A Checklist of Methodological Fallacies." Journal of Afro-American Issues, V. 4, #1, (Winter 1976), pp. 35-49.

Thoday, J.M., & Gisons, J.G. "Environmental and Genetical Contributions to Class Difference: A Model Experiment." Science, V. 167, (1970), pp. 990-992.

Thorndike, Robert L. "Concepts of Culture-Fairness." Journal of Educational Measurement, V. 8, #2, (1971), pp. 64-69.

Vance, H.R., Huelsman, C.B., & Wherry, R.J. "The Hierarchial Factor Structure of the WISC as It Relates to Disadvantaged White and Black Children." Jour. of General Psychology, V. 95, (1976), pp. 287-293.

Washington, E.D., & Teska, J.A. "Relations Between the CAT, WRAT, Stanford-Binet, and Ill. Test of Psycholinguistic Ability." Psychological Reports, V. 26, (1970) pp. 291-294.

Wellborn, E.S., Reid, W.R., & Reichard, C.L. "Effect of Examiner's Race on Test Scores of Black and White Children." Educ. and Tr. of Ment. Ret., V. 8, (1973), pp. 194-196.

Wesman, A.G. "Intelligence Testing." American Psychologist, V. 23, (1968), pp. 267-274.

Whorf, B.L. "Science and Linguistics." in E.L. Hartley and T.M. Newcomb (Eds) Readings in Social Psychology, New York: Holt, (1947).

Wikoff, R.L. "Danger: Attacks on Testing Unfair." Clinical Child Psychology Newsletter, V. 9, #3, (1970).

Williams, R.L. "From Dehumanization to Black Intellectual Genocide: A Rejoinder." Clinical Child Psych. News letter, V. 9, #3, (Fall 1970), p. 8.

Williams, R. L. "On Black Intelligence." Jour. Of Black Studies., V. 4, #1, (1973), pp. 29-40.

Williams, Robert. "Moderator Variables as Bias in Testing Black Children." Journal of Afro-American Issues, V. 3, #1, (Winter 1975), pp. 77-90.

Williams, Robert. "The BITCH-100: A Culture-Specific Test." Journal of Afro-American Issues, V. 3, #1, (1975), pp. 103-116.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Williams, Robert L. "Danger: Testing and Dehumanizing Black Children." Clinical Child Psychology Newsletter, V. 9, #1, (Spring 1970).

Zimmerman, I.L., & Woo-Sam, J.M. "Research With the Wechsler Intelligence Scale for Children 1960-1970." Psychology in the School, V. 9, (1972), pp. 232-271

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Miscellaneous

Court Case

Larry P. et al. vs. Wilson Riles, et al. #c 71-2270 RFP, U.S. District Court, Northern District of California, (June 1972).

Letters

Mercer, Jane & Lewis, June. Cover Letter for the S.O.M.P.A. Test Kit, New York: Psychological Corp., (1977).

Williams & Associates. Cover Letter for the BITCH Test Kit, Williams & Assoc., St. Louis (1972).

"Response to the Third Set of Interrogatories by S.F.U.S.D. Psychologists in the 'Larry P.' Case." Unpublished Letter, San Francisco, (May 4, 1977).

Manuscripts

Dill, J.R. " A Study of the Influence of Race of the Experimenter and Verbal Reinforcement on Creatifity Test Performance of Lower SES Black Children," Unpublished Doctoral Dissertation, New York University, Diss. Abstracts int'l., V. 32, (1972), 6071B.

Engel, Madeline. A Reconsideration of the Concept "American Lower Class: A Study of Urban Subcultures." Unpublished Doctoral Dissertation, Fordham University, New York, (1966).

Miller, J.O., & Phillipa, J.A. " A Preliminary Evaluation of the Head Start and Other Nashville Kindergartens." Unpublished Manuscript, Peabody College, Nashville, Tenn., (1966).

Reschly , D.J. & Reschly, J.E. "Validity of WISC-R Scores in Predicting Achievement and Attention for Four Sociocultural Groups," Unpublished Manuscript, Iowa State University, (1977).

Samuel, W.S., Grayson, S. Sherman, B., Soto, D.H. & Winstead, K. Motivation, Race, Social Class and I.Q.: A Follow-up Study. Paper presented at the American Psychology Assn. Convention, Wash., D.C., (Sept. 1976).

Yee, Leland Yim. Mental Abilities of Chinese Fourth Graders at Different Social Class Levels. Unpublished Master's Thesis, S.F. State Univ., SF, (1972).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Newspapers

Edson, Lee. "Jensenism, n., The Theory That I.Q. is Largely Determined by the Genes." New York Times Magazine, (8-31-69).

Pogash, Carol. "I.Q. Tests: Elementary School's Bakke Case?" S.F. Examiner, (3-3-78)

Wood, Jim. "The I.Q. Bias Battle." S.F. Chronicle Magazine, SF, (9-24-72).

Unpublished Articles

Colwell, James. "The Stanford-Binet, WISC & WAIS as They Relate to Intelligence Measurement Today." Unpublished Article, S.F.U.S.D., S.F. (1970).

Komm, Richard. "History and Problems Relating to Usage of Various Test Instruments for Minority Group Children." Unpublished Paper, S.F., (1972).

O'Neill, H.D., Jr. "Partial Validation of the Slossen Intelligence Test." Unpublished Dissertation, University of Oklahoma, Norman, Okla. (1969).

Wright, Ernst. "The Achievement of Black High School Students on the BITCH." Unpublished Study, Kern High School District, Bakersfiedl, CA. (1975).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

PLEASE NOTE:

"Black Intelligence Test of Cultural Homogeneity", copyright 1972 by R.L. Williams not microfilmed at request of author. Available for consultation at University of San Francisco Library.

UNIVERSITY MICROFILMS.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.