77-24,243

KUSE, Allan Robert, 1947-
FAMILIAL RESEMBLANCES FOR COGNITIVE ABILITIES
ESTIMATED FROM TWO TEST BATTERIES IN HAWAII.

University of Colorado at Boulder,
Ph.D., 1977
Psychology, psychobiology

Xerox University Microfilms, Ann Arbor, Michigan 48106

FAMILIAL RESEMBLANCES FOR COGNITIVE ABILITIES ESTIMATED FROM TWO TEST BATTERIES IN HAWAII

by

Allan Robert Kuse

B.S., Purdue University, 1969

M.A., University of Colorado, 1972

A thesis submitted to the Faculty of the Graduate
School of the University of Colorado in partial
fulfillment of the requirements for the degree of
Doctor of Philosophy
Department of Psychology
1977

This Thesis for the Doctor of Philosophy Degree by

Allan Robert Kuse

has been approved for the

Department of

Psychology

by

J.C. DeFries

J. C. DeFries

Daniel E Bailey

Daniel E. Bailey

Gerald E McClearn

Gerald E. McClearn

Steven G. Vandenberg

Steven G. Vandenberg

David W. Crumpacker

David W. Crumpacker

Date February 1, 1977

Kuse, Allan Robert (Ph.D., Psychology)

Familial Resemblances for Cognitive Abilities Estimated from
Two Test Batteries in Hawaii

Thesis directed by Professor John C. DeFries

Discovering determinants and correlates of cognitive functioning is a primary concern of psychological research. A large scale family study examining genetic and environmental bases of specific cognitive abilities has been conducted in Hawaii for the past five years. The Wechsler Adult Intelligence Scale (WAIS) was administered to 118 families (456 individuals) that had previously participated in this study.

Regressions of offspring on midparental scores were significant for verbal (b_{op} = .50), performance (b_{op} = .32), full scale (b_{op} = .37), and 8 of 11 WAIS subtests. Regressions were also significant for 13 of 15 tests and 5 derived factors obtained by retesting subjects with the Hawaii battery. These findings indicate that specific cognitive abilities inherent in tests of both batteries are affected by familial influences.

Comparison of the pattern of magnitudes of these familial resemblance coefficients to previous data provided evidence for the possible differential contribution of genetic factors among the various cognitive measures. Examination of familial correlations revealed no evidence of sex-linked inheritance for any ability assessed by either battery, including spatial ability; however, evidence suggestive of a maternal, environmental effect was found for 10 of 11 measures of verbal ability.

The correlation between WAIS full scale and the first principal component derived from the Hawaii battery was .73. This value is comparable to reported correlations between full scale and other standard tests of intelligence, and indicates that factor scores derived from the first component may be used as predictors of WAIS full scale for the total Hawaii data set. Further, parent-child regressions for the first component ($b_{op} = .38$) and WAIS full scale suggested that the true heritability of general cognitive ability may be lower than previously reported values.

These results provide further support to a growing body of knowledge indicating a complex array of genetic and environmental factors which, together, comprise the bases of performance on measures of cognitive ability.

This abstract is approved as to form and content. I recommend its publication.

Signed J.C. DeFries
Faculty member in charge of dissertation

ACKNOWLEDGMENTS

I would first like to express the gratitude I feel to Dr. John C. DeFries, my research advisor, for his guidance, assistance, and patience throughout my graduate career. I am indebted to the faculty, staff, and students at the Institute for Behavioral Genetics for enriching, aiding, and abetting that career. And finally, special thanks go to the faculty and staff of the Behavioral Biology Laboratory. Without these many friends, this dissertation would not be.

The results reported here were made possible by collaboration of a group of investigators (G. C. Ashton, R. C. Johnson, M. P. Mi, and M. N. Rashad at the University of Hawaii, and J. C. DeFries, G. E. McClearn, S. G. Vandenberg, and J. R. Wilson at the University of Colorado) supported by NSF grant GB 34720 and grant MH 06669 from the National Institute of Child Health and Human Development.

TABLE OF CONTENTS

CHAPTER PAGE

I. INTRODUCTION 1

II. HISTORICAL BASES 2

Wechsler Adult Intelligence Scale 4

The Inheritance of IQ 8

The Inheritance of Specific Abilities 9

Twin studies 10

Family studies 14

III. OBJECTIVES 17

Objectives Related to a Measure of General Ability. 17

Objectives Related to Familial Resemblances 18

IV. METHODS. 21

Experimental Subjects 21

Testing Procedures for Session A. 22

Cognitive Measures for Session A. 23

Testing Procedures for Session B. 24

Session B Measures. 26

V. RESULTS. 30

Description of the Sample 30

Socio-Economic Status of the Families 33

Age Corrections for WAIS Subtests 37

Correcting for Age Effects in the HFSC Battery. . . 39

Comparison of Session A Data to Previous Scores . . 41

Comparison of Session A to WAIS Scores. 46

CHAPTER PAGE

Familial Resemblances for HFSC Tests. 52

Familial Resemblances for the WAIS. 55

Other Measures of Familial Resemblance. 58

Sex-linkage. 58

Spouse correlations. 64

Sibling correlations 64

VI. DISCUSSION 66

Predictors of General Ability 66

Comparisons of Familial Resemblances. 70

Measures of Familial Correlation. 71

Familial Resemblance for IQ Measures. 73

Familial Resemblances for Specific Abilities. . . . 75

Suggestions for Further Research. 76

REFERENCES. 79

APPENDIX A: DESCRIPTIONS OF HFSC BATTERY TESTS 89

APPENDIX B: DESCRIPTIONS OF WAIS SUBTESTS. 96

APPENDIX C: ADDITIONAL TABLES REFERENCED IN THE TEXT 99

CHAPTER I

INTRODUCTION

The Hawaii Family Study of Cognition is an interdisciplinary, interuniversity study of the genetic and environmental bases of specific cognitive abilities. Between October, 1972, and August, 1976, an extensive battery of psychometric tests was administered to more than 1800 families (over 6500 individuals). Factor analyses of these data, developmental trends, and measures of familial resemblance have been previously reported (DeFries et al., 1974; Wilson et al., 1975; DeFries et al., 1976; Johnson et al., 1976).

The present dissertation is based upon analyses of data from 118 families who were retested during the summer of 1975. From theoretical considerations of the nature of cognitive tests to be reviewed in the following section, measures of general cognitive ability were derived from the test battery used in Hawaii. Concurrent validities for the measures were assessed by also administering the Wechsler Adult Intelligence Scale to the subset of families. Scores from one derived measure and both test batteries were then submitted to genetic analyses designed to test specific hypotheses concerning heritable factors in cognition.

CHAPTER II

HISTORICAL BASES

With the publication of Hereditary Genius in 1869, Francis Galton became the father of modern intelligence testing. His early operationalistic definition of natural ability included those "qualities of intellect and disposition, which urge and qualify a man to perform acts that lead to reputation [eminence]" (Galton, 1869). Although research has progressed for over a century, general intellectual ability, or intelligence, holds the dubious distinction of being the most defined, most used, most controversial, yet least understood concept in the psychological domain.

Many researchers have tried to provide a more precise definition of intelligence but, unfortunately, there is little consensus among them. Binet and Simon (1915) suggested that the essential activities of intelligence were judgement, comprehension, and reasoning. These were also embodied in Terman's (1921) definition, the ability "to carry on abstract thinking," and Burt's (1955) "innate, general, cognitive ability." Heim (1954) preferred to describe intelligent activity as consisting of "grasping the essentials in a situation and responding appropriately to them." Similarly, Humphreys (1971) defined intelligence as "the entire repertoire of acquired skills, knowledge, learning sets, and generalization tendencies considered intellectual in nature that are available at any one period of time."

Some definitions encompass all of these ideas; for instance, Wechsler (1958) stated that intelligence is "the aggregate or global capacity of the individual to act properly, to think rationally, and to deal effectively with his environment." In contrast, to paraphrase Boring (1923) and Miles (1957), who proposed and defended a definition avoiding any particulars, intelligence is that which is tested by intelligence tests.

The problem with such definitions of intelligence and the cause of their apparent diversity is that each essentially reflects the author's use of intelligence tests or his idea of intelligence as a concept. Although, in this sense, the definitions quoted above are as circular in nature as that used by Galton, the common element uniting them is an emphasis on mental activity and abstract reasoning. Another commonality was accepted by the authors, but was not evident from the definitions themselves: whatever intelligence is, a variety of psychometric tests measure some general aspect of cognitive functioning.

On the basis of such tests or batteries of tests, other authors have attempted to define intelligence, perhaps circuitously, by describing its structure through some form of multivariate analysis. Pioneered by the work of Spearman (1904), Burt (1924), and Kelley (1928), and stimulated by the research of Thurstone and Thurstone (1941), a hierarchical model of intelligence is the most pervasive and most accepted by multivariate theorists (Horn, 1972, 1976). This model assumes a large number of intercorrelated, primary mental abilities from which higher-order, general abilities can be identified. The latter are predictive of the intellectual activity to which the term intelligence has been applied.

Cattell (1971) identified two such general factors which, in turn, are themselves correlated. He further suggested that unless a broad array of primary abilities is used for derivation of the two factors, they will emerge as a single general factor. It was this general factor that Thomson (1950) described as a recurrent statistical quantity arising because every cognitive test of any complexity contains elements common to all cognitive tests; thus, if a battery is composed of a variety of tests, a general factor will necessarily emerge. Concurring with this view, Rimoldi (1951) argued that the first unrotated factor extracted in a factor analysis of cognitive tests will constitute a good estimate of general "intellective" ability.

Wechsler Adult Intelligence Scale

Although it is unsatisfactory as a definition of intelligence, the hierarchical model outlined above has been the practical basis for construction of most tests of general ability. The number and types of intelligence tests in use change from year to year as theories of intelligence evolve and psychometric techniques for test construction are refined. Buros (1965), as cited by Eber (1972), listed 82 group administered intelligence tests, 28 individually administered tests and 54 multiple aptitude batteries from which a general ability measure can be computed. Of these, the Wechsler Adult Intelligence Scale (WAIS: Wechsler, 1955) is the best known and most widely used clinical test of general intellectual ability.

Basically a revision of the Wechsler-Bellevue Scales, the WAIS is composed of both verbal and performance subtests. Wechsler chose this dichotomy because he believed that intelligence involved both abstract reasoning (verbal, spatial, and arithmetic) and the ability to handle practical situations calling for performance and manipulative skills.

Another important consideration in the selection of subtests was their utility. To be useful as measures for the entire adult population, subtests must be suitable over a wide range of ages and be reasonably balanced with regard to sex differences. They also should have "common sense appeal" in the nature and variety of tasks, so that subjects' performance would not be biased by either tedium or a lack of relevance. Finally, when all else has been considered, the group of subtests must retain the power to discriminate levels of ability among subjects (Wechsler, 1958).

Based on experience with the Wechsler-Bellevue Scales and the theoretical considerations noted above, Wechsler selected items for their statistical reliability and clinical validity. The WAIS was then standardized on a sample of 1700 individuals chosen to represent the adult population (16-64 years) in the United States and an additional 475 subjects, aged 60-75 and over (Wechsler, 1944, 1955). Since its introduction, the WAIS has been widely used for both clinical evaluation and basic research on intellectual abilities (Cohen, 1957; Shaw, 1965, 1967; Lubin, Wallis, & Paine, 1971; Zimmerman & Woo-Sam, 1973).

Wechsler (1955) reported exceptionally high reliabilities for WAIS IQ measures and substantial values for individual subtests, using split-half correlations from three age groups in the standardization sample. Later studies confirmed that the reliability of the measures was high by the method of test-retest correlation, although the results were not quite as spectacular as first assumed (Coons & Peacock, 1959; Quershi, 1968; Kangas & Bradway, 1971).

Additional research has also indicated that the WAIS possesses significant predictive and concurrent validities. Plant and Lynd (1959) reported that the WAIS predicted grades as well as the American Council on Education Psychological Examination (ACE), a measure of academic achievement. In studies of college freshmen, Giannell and Freeburne (1963) found a correlation of .84 between grade-point average and Full Scale IQ, while Olsen and Jordheim (1964) and Conry and Plant (1965) found correlations somewhat lower than this value (.58 and .62, respectively). The latter study also showed that average high school WAIS scores were consistently below those of college students, suggesting that levels of academic achievement are acceptably measured by the WAIS.

A number of studies, summarized in Table 1, indicated adequate concurrent validity for the WAIS, when this is measured by correlations with other intelligence tests. The studies included in

TABLE 1

Correlation of WAIS Scores[a] with Other IQ Measures

Test[b]	Sample	Sample Size	r_{VIQ}	r_{PIQ}	r_{FIQ}	Source
W-B I	College students	46	.87	.52	.82	Cole & Weleba (1956)
W-B II	College students	51	.77	.34	.64	Neuringer (1963)
W-B II	College students	72	.80	.66	.78	Quershi & Miller (1970)
S-B L	College students	109	-	-	.90	Giannell & Freeburne (1963)
S-B LM	Adults	48	.86	.36	.77	Kangas & Bradway (1971)
S-B LM	Gifted math students	130	.60	.09	.45	Kennedy, Willcutt & Smith (1963)
Progressive Matrices	Aged (mean age, 76)	50	.60	.49	.60	Granick (1971)

a. Correlations were computed with WAIS Verbal Scale Score (VIQ), Performance Scale Score (PIQ), and Full Scale IQ Score (FIQ).

b. W-B I = Wechsler Bellevue, Form I.
W-B II = Wechsler Bellevue, Form II.
S-B L = Stanford-Binet, Form L.
S-B LM = Stanford-Binet Revision, Form LM.

the table all used non-clinical samples, although considerable supportive data have been collected in clinical settings on samples of psychotic or mentally retarded patients (see Zimmerman & Woo-Sam, 1973). Overall, research has established the WAIS as a valid and reliable instrument for the measurement of general intellectual ability.

The Inheritance of IQ

Since the publication of Jensen's (1969) discourse on probable genetic involvement in the determination of IQ and scholastic achievement, the meaning, use, and misuse of intelligence test scores have once again been subjected to close scrutiny (e.g., Jensen, 1972; Richardson & Spears, 1972; McClelland, 1973; Kamin, 1974). A full review of this recent literature is well beyond the scope intended here, but the furor raised within the scientific community and without demands mentioning. Once again the psychometric nature of intelligence measures, their use in evaluating academic success and predicting occupational success, and their purported misuse by differentially classifying ethnic and socioeconomic groups resulted in calls for their improvement (Heim, 1970) and their elimination (Brazziel, 1969). The majority of authors have recognized the need for further research and for caution in the interpretation of results, without condemning intelligence testing *per se*.

The lengthiest debate resulting from Jensen's paper concerns the degree of genetic involvement in the determination of individual differences in IQ. Erlenmeyer-Kimling and Jarvik (1963) summarized over 50 studies of familial resemblance for IQ and concluded that the data supported a model of polygenic inheritance contributing to general cognitive ability. Reanalyses of extant data have yielded evidence for a genetic component for some investigators (Jinks & Fulker, 1970; Jencks, 1972; Eaves, 1973; Jensen, 1973; Jinks & Eaves, 1974; Rao, Morton, & Yee, 1974) but not others (Schwartz & Schwartz, 1974; Kamin, 1974). Evidence for a genetic component has also been disputed on theoretical grounds (Layzer, 1974; Lewontin, 1975; Feldman & Lewontin, 1975) and by the presentation of new data (Adams, Ghodsian, & Richardson, 1976). Only additional research will help resolve the controversy concerning the roles of genetic, environmental, and cultural factors in determining the performance of subjects on intelligence tests.

The Inheritance of Specific Abilities

Even before the IQ controversy arose, researchers had begun to explore specific cognitive abilities to help clarify the contribution of genetic factors to general cognitive ability (DeFries, Vandenberg, & McClearn, 1976). The first problem behavioral geneticists encountered was the number of abilities to consider. Thurstone (1938) in his classic monograph suggested that there are no fewer than six primary mental abilities, from which a second-order general ability measure may be extracted through factor analysis (Thurstone & Thurstone, 1941). In contrast, Guilford (1956, 1959) proposed a facet theory of intellect containing as many as 120 distinct cognitive traits. Taking a moderate view, French, Ekstrom, and Price (1963) settled on 24 primary factors, which they felt had been replicated in enough studies to be considered established primary mental abilities. However, an unwieldy 72 tests, or three reference tests for each factor, are required to define these primary abilities with any great precision.

Avoiding the dilemma of which abilities to examine first, a straightforward way to determine the relationship between specific abilities and general ability is to study the various tasks employed in standard measures of intelligence. As early as 1949, Jones discovered through factor analyses of Stanford-Binet items that a number of distinct factors were present in the test. Likewise, although WAIS subtests do not represent unique primary factors, they presumably do represent various primary abilities in different proportions; further, the wide usage enjoyed by the WAIS warrants an examination of its subtests.

Twin studies. Most researchers have used comparisons between identical (monozygotic or MZ) and fraternal (dizygotic or DZ) twins to search for evidence of an heritable component in cognitive abilities. The value of studying twins was recognized early due to the unique genetic character of MZ pairs and the developmental similarities between MZ and DZ twins (Tallman, 1928; Holzinger, 1929). Although the method does have drawbacks (Scarr, 1969; Vandenberg, 1976), it yields approximate estimates of genetic contributions to a trait. Twin research has been amply reviewed by Vandenberg (1967, 1968, 1976) and Loehlin and Nichols (1976).

Block (1968) examined the 11 subtests of the WAIS for evidence of genetic effects in one of the few twin studies using standard intelligence subtests. As an estimate of the genetic contribution on each subtest, the ratio of DZ to MZ within pair variances was used to arrive at an F-value (Vandenberg, 1967). Those subtests most highly related to verbal, space, and perceptual organization factors (Cohen, 1957) revealed the greatest degree of genetic involvement. Two subtests, Picture Arrangement and Object Assembly, showed no evidence of genetic effects. These differences among subtests, as well as differences on similar tests from other batteries, were interpreted as evidence for differential heritability of various cognitive tasks (Vandenberg, 1967, 1969). However, this is not the only possible explanation of these results. For instance, relative rankings of twin correlations among ability tests could also reflect varying test validity or reliability.

Loehlin and Nichols (1976) recently reported an analysis of data from 2164 twin pairs (1300 MZ, 864 DZ) who participated in the National Merit Scholarship Qualification Testing program (NMSQT). On the five NMSQT subtests (English Usage, Mathematics, Social Studies, Natural Science, and Vocabulary), no consistent differences in the estimated amount of genetic control were found. In other words, the degree of genetic determination (estimated as twice the difference between MZ and DZ intraclass correlations) appeared to be similar for the five tasks. Reviewing previous twin research in light of their findings, the authors observed no consistent trends in the magnitude of trait-to-trait differences between MZ and DZ correlations across studies. They interpreted these results as indicating that all cognitive abilities may tend to have moderate, roughly equal heritabilities.

Although the analyses performed by Loehlin and Nichols avoided problems of differing test validity and reliability, and possible environmental factors which might differentially affect MZ and DZ resemblances on various tests, some aspects of the study weaken their conclusion. First, NMSQT subtests do not represent specific abilities in a psychometric sense; rather, they are achievement tests. Such measures rely very heavily on verbal skills and information obtained through academic training and, with the possible exception of Vocabulary, involve specific abilities in unknown combinations. A mixture of abilities with an emphasis on those susceptible to training could easily result in similarly estimated twin correlations. In fact, each subtest might be considered a measure of general ability and would vary from the others much less than from tests of specific traits, relatively unrelated to skills acquired in school. The consistency of the median MZ and DZ correlations among NMSQT subtests, shown in Table 2, supports this interpretation. Vandenberg (1969) has shown that even among tests of spatial ability, magnitudes of MZ (and DZ) correlations widely vary.

The Vandenberg study is equally relevant to Loehlin and Nichols' analytical review of previous twin research. Since their analyses included quite different measures for each specific ability, differences among tests purportedly measuring the same

TABLE 2

Median Intraclass Correlations of Five NMSQT Subtests[a]

Subtest	Identical Twins	Fraternal Twins	Difference
English Usage	.72	.52	.20
Mathematics	.71	.48	.23
Social Studies	.73	.52	.21
Natural Science	.65	.48	.17
Vocabulary	.87	.62	.25

a. Adapted from Loehlin and Nichols (1976: Table 4-4, p. 33).

cognitive trait could easily have caused the observed lack of consistency. Because of small samples, most of the studies reviewed by Loehlin and Nichols were also prone to considerable errors in estimation of twin correlations. Fairly large samples are required to reduce confidence intervals of correlations to reasonable levels. With regard to specific cognitive traits, this statistical necessity makes an adequate test of Loehlin and Nichols' hypothesis prohibitively expensive using the twin method. However, analyses based on resemblances between family members other than twins offer an alternative source of information.

Family studies. Data are available from only one family study of specific cognitive traits, the Hawaii Family Study of Cognition (HFSC), which was large enough to be useful in evaluating the hypothesis of Loehlin and Nichols. Regressions of mid-child on mid-parent scores for 15 cognitive tests in the HFSC were reported by DeFries and his colleagues (1976) for two ethnic groups (6 of 15 were reference tests for specific cognitive factors identified by French et al., 1963). A Spearman rank correlation between sets of regressions for Americans of European Ancestry (AEA) and Americans of Japanese Ancestry (AJA) families was .77 ($p < .01$). In addition, regressions for three tests (Vocabulary, Pedigrees, and Progressive Matrices) were significantly greater than those for tests of visual memory and maze tracing in the AEA sample. Although much of the same trend was evident in the AJA sample, the comparisons did not reach significance, perhaps due to a much smaller number of families (244 compared to 739 AEA families). These results tend to support Vandenberg's (1967) suggestion that different mental abilities are determined to different degrees by hereditary factors. Further, since the regressions had been corrected for test reliability and the relative pattern of coefficients was highly consistent despite possible cultural differences, the results are probably quite robust.

Although using WAIS subtests rather than tests of specific abilities, a second family study (Williams, 1975) provides a comparison for the twin data of Block (1968). Differences between the two studies are notable. The family study used regressions of sons' scores on midparental values as a measure of familial resemblance, while Block reported data based on MZ and DZ twin differences. The former study employed a Canadian sample, and the collection of twin data was centered in Louisville, Kentucky. Finally, sons in the family study were given the Wechsler Intelligence Scale for Children, rather than the WAIS. Although the two batteries contain different subtests, both were used in the family study because they measure similar abilities in different age groups (Williams, 1975).

For ease of comparison, results obtained by Williams and Block are shown in Table 3. Overall, the Arithmetic, Vocabulary, and Digit Symbol subtests most clearly indicated a genetic component. On the other hand, Object Assembly and Picture Completion showed no such evidence. A rank correlation of .65 ($p < .05$) between the sets of tabled values also suggested that relative sizes of estimated genetic contributions across the 11 subtests were consistent. Thus, taken together, the results of Block and Williams support a hypothesis of differential genetic determination of various cognitive traits.

TABLE 3

Resemblance Among Relatives on WAIS Subtests

	60 MZ and 60 DZ Pairs		55 Mid-parent and Son Pairs	
Information	3.88	***	.25	*
Comprehension	2.55	**	.26	*
Arithmetic	2.78	***	.45	**
Similarities	1.81	*	.36	*
Digit Span	1.53	*	.13	
Vocabulary	3.14	***	.53	**
Digit Symbol	2.06	**	.56	**
Picture Completion	1.50		-.12	
Block Design	2.35	**	.34	*
Picture Arrangement	1.74	*	.07	
Object Assembly	1.36		.02	
	F-values[a] from Block (1968)		Parent-child regressions[b] from Williams (1975)	

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

a. In the Block study, significance was estimated by an F-test computed as the ratio of DZ to MZ, within pair variances.

b. In the Williams study, significance was determined from the regression of sons' on midparental values.

CHAPTER III

OBJECTIVES

In a large research program entitled the Hawaii Family Study of Cognition (HFSC), collaborators from the University of Hawaii and the University of Colorado have attempted to identify environmental and genetic correlates of performance on tests of specific cognitive abilities. The primary objectives of the present research were twofold: first, to derive a measure of general cognitive ability from the HFSC battery and assess its validity by comparison to a widely accepted measure, WAIS Full Scale IQ; secondly, to obtain estimates of familial resemblance for WAIS subtests and HFSC tests and derived scores.

Objectives Related to a Measure of General Ability

(1) Subjects were a subset of families who had previously participated in the HFSC. Sampling was based solely on the date of original testing; therefore, it was reasonable to assume that subjects would be representative of the entire HFSC sample. To the extent this assumption was established, results could be generalized to the larger sample.

(2) HFSC battery data used in the study were obtained by retesting the subset of families. If these scores correlated well with subjects' original scores, it would establish that discrimina-

tion among ability levels was maintained at retesting. Also, if the retest factor structure was congruent with that obtained in the HFSC, it would demonstrate that the pattern of test intercorrelations was also unchanged. Again, the lack of these findings would affect any inference to the HFSC sample.

(3) On theoretical bases, the first principal component factor computed from the 15 HFSC tests should be a good predictor of general cognitive ability. Specifically, factor scores for individuals derived from the first component should correlate well with WAIS Full Scale IQ scores.

(4) As a corollary to the second and third points, scores from spatial and verbal rotated factors extracted from HFSC tests should have high correlations with WAIS Performance and Verbal Scales, respectively. These associations were predicted from similarities between some HFSC tests and WAIS subtests.

Objectives Related to Familial Resemblances

(1) Techniques used to estimate familial resemblance were derived from the empirically supported theory of quantitative genetics. However, various methods used to obtain best estimates from actual data often yield strikingly dissimilar results. To gauge possible estimation bias introduced because of varying numbers of offspring in sample families, regressions of offspring on mid-parental values were computed with three theoretically similar, yet statistically different methods: a) only one child was used from each family, losing information available from additional progeny; b) the average score of offspring was used, allowing each family the

same weight but ignoring within-sibship variance; and c) each offspring was regressed on the midparental score, weighting the progeny means in proportion to the number tested in each family. If concordant estimates were obtained among the three methods, it would be unlikely that the use of different estimates would cause appreciable bias across studies.

(2) Effects due to age have been described for WAIS subtests (Wechsler, 1958) and HFSC tests (Wilson et al., 1975). Since the familial resemblance measures used in this study required comparisons of individuals of quite different ages, it was necessary to adjust for the relationship between age and test performance. Corrections for the effect of age have been developed for WAIS scores (Wechsler, 1955) and the HFSC battery (DeFries et al., 1976). These corrections were applied to the scores prior to any analyses.

(3) DeFries et al. (1976) reported estimates of familial resemblance separately for AEA and AJA subjects[1] on cognitive tasks in the HFSC. Since the same set of tests was used as part of this study, coefficients of resemblance were computed for the total sample and separately for AEA subjects, the only large, ethnically homogeneous group retested. Estimates of resemblance computed from this sample should be comparable to those reported by DeFries and his co-authors.

(4) Finally, the pattern of familial resemblance coefficients for WAIS subtests was examined for evidence of differential heredi-

---

[1] Americans of European Ancestry and Americans of Japanese Ancestry.

tary influence among the various cognitive tasks. If the relative magnitudes of resemblances for subtests were consistent with those reported by Block (1968) and Williams (1975), it would strongly suggest systematic, predictable differences in familial resemblance among the subtests. Further, if some subtests showed no resemblance among family members across studies, a strong argument could be made for the absence of any heritable component for these tasks.

CHAPTER IV

METHODS

Experimental Subjects

The Hawaii Family Study of Cognition (HFSC) was initiated in 1972. Family participation was solicited by radio and newspaper announcements, letters sent to parents of students, circulars displayed in health clinics, contacting clubs and organizations, and referral by previously participating families. As a result of this effort, 1819 families were recruited and tested at the Behavioral Biology Laboratory between October, 1972, and August, 1976.

A group of 108 families was retested with the entire HFSC battery during the summer of 1974 in an effort to assess reliabilities and developmental changes for the measures.

The data analyses reported here include 118 families (456 individuals) who volunteered to attend two additional testing sessions during the summer of 1975. These will be referred to as Session A and Session B. Both parents and 64 offspring from 61 of the families who had been retested the year before were part of this sample.

Although representation of the entire HFSC sample was desirable, there were three criteria followed in the solicitation of families. First, it was required that at least nine months had elapsed since initial participation in the HFSC. Secondly, no family was solicited in which a member had had missing cognitive data for the HFSC battery. Missing data were considered a possible indication of unwillingness or inability to perform the requisite tasks. Finally, families were encouraged to include offspring who had not been subjects previously, in an attempt to increase the average family size of the sample. Of 165 families solicited, 72% participated in the additional research.

Testing Procedure for Session A

At the first testing session, Session A, groups of from 10 to 25 families were retested with the entire battery of HFSC cognitive measures. For the convenience of families, testing was conducted six times during a month: a Saturday morning and afternoon, a Sunday morning and afternoon, and two Wednesday evenings. Excluding the collection of blood and saliva samples, the procedures followed were nearly identical to those used by HFSC (Wilson et al., 1975).

As subjects arrived at the testing site, those over 18 signed a statement indicating their willingness to participate in the research. In addition, parents were requested to sign a form consenting to the testing of their minor children. Each subject received a set of gummed labels containing a unique code number by which all testing materials used during the sessions were identified. Doughnuts, coffee, and fruit punch were available for refreshment at day sessions, and box-dinners were provided at evening sessions.

After all subjects had completed the sign-in procedures, families were seated in the testing auditorium according to a pre-arranged plan which reduced the opportunity to talk or collaborate. A staff member then delivered a brief welcoming address and familiarized the subjects with test materials and procedures. To assure comparability of the various testing occasions, the battery of measures was administered using an Optisonics Sound-o-matic III Programmer-Recorder,[1] which both presented instructions with synchronized 35 mm slides and controlled timing for the cognitive tests. The machine was programmed to pause after each set of instructions to permit clarification by the test monitor whenever subjects had questions.

Subjects were given a ten minute recess after an hour of testing. During the second hour, another 40 minutes of cognitive tasks and a personal history questionnaire were administered. As subjects finished the questionnaire, they were asked to return the test materials before leaving the auditorium. Each family received compensation upon completing all procedures ($40 plus $10 for each offspring tested).

Cognitive Measures for Session A

The 15 tests used in the HFSC and at Session A were selected on the basis of a pilot study of 46 measures (Vandenberg, Meredith, & Kuse, unpublished). In order of administration, each test, its

---

[1] Sound-o-matic is a trademark of the Optisonics Corporation, Montgomeryville Industrial Park, Montgomeryville, Pa., 18936.

source, and published reliability estimate are listed in Table 4. A short description of each measure is included in Appendix A.

Because of the large amount of material in the battery, a number of steps were taken to minimize the time required for testing, and, as much as possible, to assure standardization of procedures for subjects tested at different times. The use of tape recordings for test presentation has already been mentioned. To expedite testing further, the 15 tests and personal history questionnaire were reproduced as a spiral-bound booklet and, to facilitate monitoring, successive tests were printed on different colored pages. Finally, due to the varied nature of the materials, subjects were instructed to write answers directly in the booklets, thus avoiding problems associated with separate answer sheets.

Testing Procedure for Session B

The second testing occasion was procedurally different from Session A since the tasks included individually administered tests, group administered timed-tests, and self-administered inventories. Because of limitations in facilities and available staff, only two families (up to eight persons) could be tested at one time. Sessions were scheduled seven times per week for nine weeks: on Tuesday, Wednesday, and Thursday evenings, and morning and afternoon on both Saturday and Sunday. As at Session A, each family was asked to sign a consent statement, and a staff member explained the procedures to be followed throughout the session.

TABLE 4

Session A Cognitive Tests, Test Times, and Reliabilities[a]

Test	Test Time	Reliability	Abbreviation
Vocabulary, Primary Mental Abilities (PMA)	3 min.	0.96 (PUBL)	VOC
Visual Memory (immediate)	1-min. exposure/ 1-min. recall	0.58 (KR-20)	VM1
Things (a fluency test)	2 parts/3 min. each	0.74 (CRα)	TH
Sheppard-Metzler Mental Rotations (modified for group testing by Vandenberg)	10 min.	0.88 (KR-20)	MR
Subtraction & Multiplication	2 parts/2 min. each	0.96 (CRα)	SAM
Elithorn Mazes ("lines & dots"), shortened form	5 min.	0.89 (PUBL)	LAD
Word Beginnings & Endings Educational Testing Service (ETS)	2 parts/3min. each	0.71 (CRα)	WBE
ETS Card Rotations	2 parts/3 min. each	0.88 (CRα)	CR
Visual Memory (delayed recall)	1 min.	0.62 (KR-20)	VM2
PMA Pedigrees (a reasoning test)	4 min.	0.72 (PUBL)	PED
ETS Hidden Patterns	2 parts/2 min. each	0.92 (CRα)	HP
Paper Form Board	3 min.	0.84 (KR-20)	PFB
ETS Number Comparisons	2 parts/1.5 min. each	0.81 (CRα)	NC
Whiteman Test of Social Perception	10 min.	0.69 (KR-20)	SPV
Raven's Progressive Matrices, shortened form	20 min.	0.86 (KR-20)	PMS

a. PUBL = from test manual; KR-20 = Kuder-Richardson Formula 20; CRα = Composite Reliability Coefficient (Cronbach, 1951; Lord & Novick, 1968; from Wilson et al. (1975).

Members of one family were escorted to private testing rooms and given individually administered test batteries by trained examiners. Since the adjoining rooms were not entirely sound-proofed, testing was begun with alternate batteries to avoid subjects overhearing correct responses for a test being given concurrently in the next room. A short break was permitted after an hour or a half hour of testing, depending on the battery administered first.

Meanwhile, the second family was administered an hour-long battery of cognitive and information-processing measures. After a short break, the family completed two self-administered personality inventories. After finishing one set of procedures, the families were given the remaining group of tests, so the entire session lasted about three hours. Upon completing all procedures, families received compensation as at Session A.

Session B Measures

Tests from Session B are presented in order of administration, with the time required for each, in Table 5. Of the measures collected, only the Wechsler Adult Intelligence Scale (WAIS: Wechsler, 1958) was included in the present analyses.

The WAIS consists of a battery of 11 subtests. Six are combined to yield the Verbal Scale (VIQ); the remaining five comprise the Performance Scale (PIQ); and all 11 subtests are used to derive the Full Scale score (FIQ). The names of subtests and their published reliabilities are given in Table 6, and brief descriptions of

TABLE 5

The Measures and Administration Times for Session B

Measure		Time
Individually Administered Tests		
Wechsler Adult Intelligence Scale (WAIS)		60 minutes (approx.)
Time Interval Estimation		5 minutes
(short break)		5 minutes
Peabody Individual Achievement (PIAT) Recognition, Comprehension, and Spelling subtests		25 minutes
	TOTAL	1 hour 35 minutes
Individually Administered Tests (alternate order)		
PIAT subtests		25 minutes
Time Interval Estimation		5 minutes
(short break)		5 minutes
WAIS battery		60 minutes (approx.)
	TOTAL	1 hour 35 minutes
Group Administered Tests		
Visual Information Processing (3 parts)		15 minutes
Auditory Information Processing (3 parts)		15 minutes
Perception of Simultaneity		8 minutes
Figure Memory I		2-1/2 minutes
Colorado Perceptual Speed Test (3 parts)		5 minutes
Figure Memory II		1-1/2 minutes
Identical Blocks		15 minutes
(short break)		5 minutes
Comrey Personality Scales		20 minutes
Eysenck Personality Questionnaire		10 minutes
	TOTAL	1 hour 37 minutes

TABLE 6

Abbreviations and Reliabilities for WAIS Subtests

Subtest	Abbreviation	Reliability[a] 18-19 Yrs Old	45-54 Yrs Old
Information	I	.91	.92
Comprehension	C	.79	.79
Arithmetic	A	.79	.86
Similarities	S	.87	.85
Digit Span	D	.71	.66
Vocabulary	V	.94	.96
Verbal Scale (I+C+A+S+D+V)	VIQ	.96	.96
Digit Symbol	DS	.92	---
Picture Completion	PC	.82	.83
Block Design	BD	.86	.82
Picture Arrangement	PA	.66	.74
Object Assembly	OA	.65	.71
Performance Scale (DS+PC+BD+PA+OA)	PIQ	.93	.94
Full Scale (VIQ+PIQ)	FIQ	.97	.97

a. Split-half reliabilities reported by Wechsler (1958) for two age groups that are well represented in the present sample.

the subtests are presented in Appendix B. Since many subtests either have no time limit or times vary from subject to subject, administration of the battery required from 50 minutes to an hour and 15 minutes. However, most subjects completed the WAIS in about an hour.

CHAPTER V

RESULTS

Description of the Sample

Although the families who volunteered to participate were drawn from the HFSC sample, none of the available demographic data was used in soliciting participation. Thus, with regard to family size, ethnicity, age, and socio-economic status (SES), the composition of this sample and its representation of the larger group were unknown until after testing had been completed. One difference in the subsample was intended, however, since families were encouraged to include offspring not previously tested.

This effort to increase the average family size over that of the HFSC was quite successful; an additional 27 male and 25 female offspring were recruited, resulting in a mean family size of 3.9 with a modal value of four members. In contrast, the HFSC had a mean family size of 3.5 members and a mode of three members (HFSC staff report, 1976). The distribution of the sample by family size is presented in Table 7, which also gives the number of offspring of each sex for every family size. In all, there were 111 male and 109 female offspring equally distributed by sex among different sized families ($\chi^2_{(3)} = .6$, $p > .8$).

TABLE 7

Distribution of Sample by Family Size and Sex of Offspring

Family Size	Number of Families	Number of Male Offspring	Number of Female Offspring
3	42	21	22
4	56	57	55
5	14	23	19
6	6	11	13
Total	118	111	109

The ethnicity of each subject was determined from the reported ethnicity of that subject's parents. Broken down into four major groups, the ethnic composition of the sample is given in Table 8. Americans of European Ancestry (AEA) comprised 43% of the families; Americans of Japanese Ancestry (AJA), 31%; and Americans of Chinese Ancestry (ACA), 13%. The remainder of the sample (14%) was ethnically heterogeneous, containing cross-ethnic marriages and individuals of mixed European, Oriental, Hawaiian or Filipino backgrounds. There was no evidence that the defined ethnic groups differed in family size ($X^2_{(3)} = 1.36$, $p > .2$) or the sex of offspring ($X^2_{(3)} = .8$, $p > .3$), so it may be concluded that no inadvertant sampling bias occurred with regard to these three factors. However, the ethnic composition differed significantly from that in the total group of HFSC families ($X^2_{(3)} = 17.2$, $p < .005$). The HFSC assortment was: AEA, 52.5%; AJA, 20.1%; ACA, 5.5%;

TABLE 8

Ethnic Composition of Sample Families

Ethnic[a] Group	Number of Families	Percent of Families	Number of Individuals	Percent of Individuals
AEA	51	43.2	200	43.9
AJA	37	31.4	135	29.6
ACA	13	11.0	53	11.6
Other	17	14.4	68	14.9
Total	118		456	

a. Americans of European Ancestry (AEA)
Americans of Japanese Ancestry (AJA)
Americans of Chinese Ancestry (ACA)

and the remainder, 21.9% (HFSC staff report, 1976). As compared to the larger group, the sample had an over-representation of oriental families, both AJA (+11%) and ACA (+5.5%), and an under-representation of AEA (-9%) and others (-7.5%). This discrepancy may have resulted from a bias in the solicitation procedure or a greater willingness on the part of the oriental families to participate in the additional research.

The sample parents ranged from 34 to 62 years of age, and the offspring from 14 to 27 years. Table 9 shows the distribution of subjects over the seven age groups used in the standardization of the WAIS. An eighth group was included for 14 and 15 year olds because the sample extended below the range of the WAIS standardization. The age distribution for males did not differ significantly from that of females ($\chi^2_{(7)} = 10.3$, $p > .15$). Only in the smallest age category, containing nine offspring 25 to 27 years old and one 34 year old mother, was there an overlap between the parental and offspring generations.

TABLE 9

Age Distribution of the Sample

Age Group	Number of Males	Percent of Males	Number of Females	Percent of Females	Number of Subjects	Percent
14-15	22	9.6	18	7.9	40	8.8
16-17	37	16.2	48	21.1	85	18.6
18-19	26	11.4	24	10.6	50	11.0
20-24	22	9.6	14	6.2	36	7.9
25-34	4	1.7	6	2.6	10	2.2
35-44	35	15.3	52	22.9	87	19.1
45-54	67	29.3	56	24.7	123	27.0
55-64	16	7.0	9	4.0	25	5.5
Total	229		227		456	

Socio-Economic Status of the Families

The socio-economic status (SES) of the families was determined in two ways. First, the occupations of the husbands were classified according to the Duncan rating system, which was derived to include both the income level associated with a specific occupation and the status accorded the occupation by a random sampling of individuals (Reiss, 1961). As a second measure, the educational levels of both parents in each family were examined.

In Figure 1 the occupational ratings of husbands are presented for the AEA and total samples. The distribution for all husbands, with the exception of eight who reported no occupation, had a range of 54 to 93 with median of 74 and mode of 83. A Duncan rating of 54 is assigned to a truck driver, while a dentist receives a rating of 93. The median rating was that for a trained technician, and the modal value was indicative of a career in engineering (see Appendix B in Reiss, 1961). Excluding four without occupational ratings, comparable statistics for AEA husbands were: range, 56 to 85; median, 74; and mode, 72.

In order to test for a discrepancy in occupational attainment between AEA and non-AEA husbands (total less AEA), the three lowest categories of Figure 1, and the three highest, had to be collapsed because of low frequencies at the extremes of the distributions. A chi-square test showed that the distribution of occupational status among AEA's was not significantly different from that of non-AEA husbands ($\chi^2_{(3)} = 3.3$, $p > .3$). Three husbands among the families of mixed ethnicity were AEA and have been included in that group for this comparison.

Since the reported occupations of the husbands were more often than not of a technical or professional nature, it should be expected that their educational attainment was also high. Table 10 shows the number of individuals in the AEA and total samples, partitioned by sex and generation, who had completed various

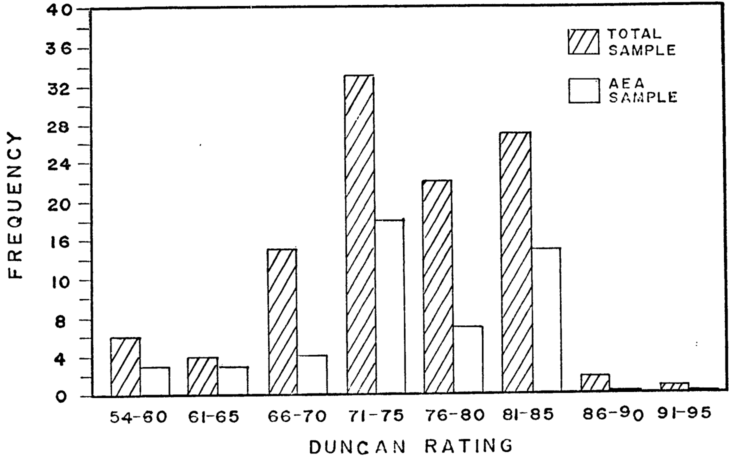

FIGURE 1. Distributions of occupational ratings for husbands in the total and AEA sample. Four AEA husbands and eight in the total sample did not report their occupations.

TABLE 10

Educational Attainment for the AEA and Total Samples[a,b]

		AEA Sample				Total Sample			
		Husbands	Wives	Sons	Daughters	Husbands	Wives	Sons	Daughters
Junior high	% =	1.9	0.0	49.1	59.0	1.7	0.8	49.0	57.8
	N =	(1)	(0)	(25)	(23)	(2)	(1)	(50)	(56)
High school	% =	7.4	26.8	29.4	28.2	11.0	22.9	28.4	28.8
	N =	(4)	(15)	(15)	(11)	(13)	(27)	(29)	(28)
Some college or technical/ business school	% =	33.3	44.6	17.6	12.9	32.2	43.2	18.6	9.3
	N =	(18)	(25)	(9)	(4)	(38)	(51)	(19)	(9)
Obtained baccalaureate degree	% =	20.4	10.7	3.9	2.6	31.4	10.2	3.9	4.1
	N =	(11)	(6)	(2)	(1)	(37)	(12)	(4)	(4)
College beyond baccalaureate	% =	37.0	17.8	0.0	0.0	23.7	22.9	0.0	0.0
	N =	(20)	(10)	(0)	(0)	(28)	(27)	(0)	(0)

a. This information was not available for 21 offspring (9 sons and 12 daughters), 8 of which were AEA.

b. Reported values indicate the number of parents who had completed a specified level of education and the number of offspring who were either enrolled in or had quit school at that level.

educational levels. Among the total sample, all but two husbands and one wife had completed high school; 87% of the husbands and 76% of the wives reported some training beyond the high school level; and 24% of the husbands and 23% of the wives reported some training beyond the baccalaureate. In short, the sample of parents was highly educated as a whole.

While there was no significant difference between AEA wives and non-AEA wives in the distribution of educational achievement ($\chi^2_{(3)} = 1.7, p > .6$), AEA husbands differed from their non-AEA counterparts ($\chi^2_{(3)} = 8.8, p < .01$).[1] Although 53% of the non-AEA husbands, as compared to 57% for AEA's, had obtained at least one college degree, the discrepancy was primarily due to fewer non-AEA husbands (12.5% as compared to 37% AEA) continuing in school beyond the baccalaureate. Since schooling is compulsory for minors, it was not surprising that AEA offspring did not differ from their cohorts in the level of educational attainment ($\chi^2_{(3)} = .2, p > .95$).

Judging from both occupational status and amount of education, the sample parents were more affluent and more highly educated than the general population. Thus, it appears that subjects were drawn from the middle to upper-middle classes of American society.

Age Corrections for WAIS Subtests

Means and standard deviations for total sample WAIS scores are presented in Table 11. Subjects performed at a high level on

---

[1] Three AEA husbands and five AEA wives from mixed ethnicity families were included with the AEA sample for these comparisons.

TABLE 11

Means and Standard Deviations of Scaled Scores for the WAIS

	Verbal Scales		n=456		Performance Scales	
Subtest	Mean	S.D.		Subtest	Mean	S.D.
I	11.3	2.8		DS	12.4	2.7
C	11.4	3.4		PC	11.1	2.2
A	11.4	2.8		BD	12.0	2.8
S	11.8	2.2		PA	10.6	2.4
D	11.0	2.9		OA	10.7	2.6
V	11.6	2.8				
VIQ	112.8	11.1		PIQ	114.8	10.3
			Full Scale	FIQ	114.5	9.9

nearly all subtests and VIQ, PIQ, and FIQ, averaging almost a full standard deviation above the WAIS standardization sample on these composite measures. As expected (Jensen, 1969), the AEA sample performed slightly better on many of the verbal subtests, while the primarily oriental, non-AEA subjects did better on performance subtests. There was no difference between these groups, however, on the Full Scale measure ($F_{(1,454)} = .59$, $p > .6$).

Wechsler found significant variation among age groups in the WAIS standardization sample, so he provided tables of age-scaled scores in the WAIS manual (Wechsler, 1955), in addition to scaled score equivalents of raw scores. Similar differences were found in this sample, so it seemed desirable to use the same corrections for age differences as Wechsler. His age-scale tables were incorporated into a Fortran computer program to correct for differences among the groups, as defined in Table 9. Because the WAIS had been standardized only for subjects who were 16 years or older, age-adjusted scores for the 14-15 year group were derived by first transforming raw scores to yield the same mean and standard deviation as the 16-17 year group, and then using the age-scaling tables from the WAIS manual computed for the older group. The method was deemed appropriate because distributional properties on the subtests, other than means and standard deviations, were similar for the two age categories. To determine the efficacy of the adjustment, corrected scores for the composite measures and the 11 subtests were submitted to one-way ANOVA's across age categories, the results of which are summarized in Table 12. The analyses indicated that significant group differences remained for seven subtests and the composite scores; and further, the linear correlation with age group was also highly significant for the composites and four subtests.

To remedy these differences among groups, a second set of age corrections was applied to the WAIS data. Raw scores for every subtest and composite score were standardized within each of the eight age categories, which effectively eliminated the among-group age differences.

Correcting for Age Effects in the HFSC Battery

Scores from the battery of 15 tests were adjusted for age differences with norms derived from a sample of 5028 individuals.

TABLE 12

Efficacy of Age-Scaled Score Equivalents[a] For the WAIS Subtests and Composite Scores

Subtest	F-Value For Difference Among Age Groups	Probability of Difference	Correlation With Age Group	Significance of Linear Component	Significance of Non-Linear Components
Information	10.32	<.001	.26	<.001	<.001
Comprehension	6.94	<.001	.24	<.001	<.04
Arithmetic	1.25	N.S.	-.07	N.S.	N.S.
Similarities	0.97	N.S.	.01	N.S.	N.S.
Digit Span	0.79	N.S.	.05	N.S.	N.S.
Vocabulary	3.00	<.01	.02	N.S.	<.01
Digit Symbol	8.25	<.001	.30	<.001	N.S.
Picture Completion	2.51	<.03	.08	N.S.	<.05
Block Design	2.46	<.04	-.00	N.S.	<.05
Picture Arrangement	7.83	<.001	.28	<.001	N.S.
Object Assembly	1.19	N.S.	.09	N.S.	N.S.
Verbal Scale	4.55	<.001	.18	<.001	<.05
Performance Scale	4.98	<.001	.20	<.001	N.S.
Full Scale	5.54	<.001	.21	<.001	<.05

---

[a] These were obtained from the tables of "IQ Equivalents of Sums of Scales Scores" and "Age-Scaled Score Equivalents of Raw Scores" in the WAIS Manual (Wechsler, 1955, pp. 77-91 and pp. 101-107).

Although similar to the age-banding technique used for the WAIS scores, the HFSC standardization employed 34 age groups defined as follows: year intervals for ages 14-20; 21-22; 23-27; 28-33; yearly intervals from 34-55; 56-57; and ages 58-72. So that no category would have a frequency of under 50, the width of each interval was determined from the number of individuals who would be included. This technique resulted in a median group size of 121 with age group frequencies ranging from a low of 51 to a high of 476 individuals. Within each interval, scores were standardized for all 15 tests, eliminating both linear and non-linear differences among the 34 age groups. Utilizing the HFSC norms to adjust the data also allowed direct comparisons of subjects' scores from Session A with their original scores obtained on the HFSC battery.

Comparison of Session A Data to Previous Scores

All but 21 offspring in the sample had previously been administered the HFSC battery as a part of the larger study. For 57 of the families (221 individuals), Session A was the second administration; in the remaining 61 families, 28 offspring had participated only once before, and 186 subjects had been tested on two prior occasions. On the average, about 15 months elapsed between testing sessions for the 57 families, and about 12 months for subjects tested twice before. Although this lack of constancy in Session A confounds many comparisons that can be made with the original test scores, a few are both necessary and informative.

First, it was reasonable to assume that scores of subjects would show some improvement in the retest data. To test this assumption, age-corrected scores of subjects when initially tested were subtracted from age-corrected scores obtained at Session A. Since the same adjustment for age effects had been applied to both data sets, and this method resulted in standardized scores for the HFSC sample, the difference scores were scaled in standard deviation units.

Means and standard deviations for both the Session A tests and the differences scores are presented in Table 13. On 11 tests--VOC, VMI, TH, MR, LAD, WBE, CR, PED, HP, SPV and PMS--significant increases were found ranging from .09 (VOC) to .65 (PED) standard deviations; on three tests--SAM, VMD and PFB--no change was evident; and for one test, NC, a significant decrement was observed. Three of five tests measuring aspects of spatial ability--MR, LAD and HP--showed large increases, which may have resulted from a clearer understanding of the required task or as a consequence of the practice afforded by prior testing. The latter explanation, though, is less plausible when the length of the intervening period is considered. Another interesting observation was that, of two measures of numerical speed and accuracy, NC showed a significant decrement and SAM showed no change between testing occasions.

Table 13 also presents test-retest correlations, an estimate of reliability for the measures. With the exception of that for PED, these estimates of reliability are lower than the published values reported in Table 4; however, the relative magnitudes of the

TABLE 13

Means and Standard Deviations for Session A Tests, Differences between Session A and First Testing and Test-Retest Reliabilities

Test	Session A		Difference			Reliability
	Mean Session A Score	S.D. of Session A Score	Mean Difference Score	S.D. of Difference Scores	T-Value	Test-Retest Correlation
VOC	.31	0.89	.09	0.43	4.50*	.77*
VMI	.10	0.97	.13	0.95	2.77*	.45*
TH	.41	1.06	.29	0.72	8.19*	.66*
MR	.56	1.01	.42	0.66	13.11*	.70*
SAM	.22	1.08	-.02	0.42	0.94	.83*
LAD	.62	1.04	.58	1.03	11.44*	.40*
WBE	.37	1.00	.32	0.68	9.59*	.68*
CR	.36	1.02	.16	0.66	5.07*	.72*
VMD	.05	0.98	.06	0.92	1.42	.51*
PED	.76	0.98	.65	0.59	22.57*	.73*
HP	.71	1.04	.52	0.84	12.60*	.59*
PFB	.16	0.96	-.06	0.73	1.66	.60*
NC	-.04	0.99	-.23	0.62	7.63*	.69*
SPV	.44	0.90	.35	0.74	9.67*	.51*
PMS	.46	0.89	.28	0.62	9.45*	.64*

*$p < .001$.

coefficients were significantly related (Spearman Rank Correlation = .57, $p < .05$). The exact source of error variance causing lower reliability estimates was not determined, although it most plausibly could be attributed to changes in subjects during the lengthy intervals between testing and retesting.

The final comparison involved factors extracted from the HFSC data (DeFries et al., 1974) and factors from Session A. Factor scores from the analyses of Wilson et al. (1975), which confirmed the findings of the 1974 report, were available for all but 21 of the subjects. Employing the same procedure as these studies, the data from Session A were subjected to principal component analysis and subsequent varimax rotation. The number of factors retained for rotation was determined by the number of eigenvalues exceeding one (Kaiser, 1960). Table 1-1 in Appendix C contains the data correlation matrix for the total sample, and Tables 1-2 and 1-3 present the unrotated and varimax loading matrices, respectively. Similarly, Tables 1-4, 1-5, and 1-6 in Appendix C present the correlations, unrotated, and rotated factor loadings for the AEA sample. After rotation, four interpretable factors emerged: SPATIAL; VERBAL; SPEED, a numeric speed and accuracy factor; and MEMORY, a factor defined by VMI and VMD. The four rotated factors and the first principal component (FIRST PCF), interpreted as a measure of general cognitive ability (Rimoldi, 1951), were strikingly similar to the factors reported for the HFSC sample. This was confirmed by coefficients of congruence computed between factor loadings from Session A and those of HFSC: .994 for SPATIAL; .991

for VERBAL; .985 for SPEED; .979 for MEMORY; and .998 for FIRST PCF. By direct solution, factor scores were produced for the five measures. The intercorrelations of the two sets of factor scores, shown in Table 14, verified that the four cognitive functions identified by the rotated factors in both data sets were similar in nature and mutually independent. Diagonal elements in the matrix can be interpreted as test-retest measures of reliability. As such, these values were remarkably high, considering that the average interval between testing occasions was more than one year. Further, very low correlations among the orthogonal factors provided convincing evidence of the stability of the factor structure obtained through principal component analyses of the HFSC tests.

TABLE 14

Intercorrelations of Factor Scores Obtained From First Testing and Session A

Factors From HFSC	Session A Factors SPATIAL	VERBAL	SPEED	MEMORY	FIRST PCF
SPATIAL	.84*	.00	.02	-.02	.54*
VERBAL	-.04	.83*	-.02	-.02	.45*
SPEED	.05	.01	.86*	.03	.43*
MEMORY	-.01	.01	.04	.60*	.14*
FIRST PCF	.56*	.53*	.37*	.13*	.87*

*$p < .05$

Comparison of Session A to WAIS Scores

Correlations between the HFSC battery administered at Session A and WAIS subtests and composite scale scores are reported in Table 2-1 of Appendix C. These values were not corrected for measurement unreliability or restrictions of sample range. Although each of these procedures is psychometrically valid, uncorrected values are reported because they represent conservative estimates of association for the HFSC as a whole.

The intercorrelation of the verbal and performance composite scores ($r = .52$, $p < .001$) demonstrated that VIQ and PIQ both tap a common set of cognitive abilities to a significant degree. Another indication of this overlap is found in Table 2-1. The measures in the HFSC battery were selected to represent an array of specific cognitive abilities, yet all of the correlations between these 15 tests and the WAIS composite scores were highly significant ($p < .001$). This high degree of association strongly suggested that the cognitive domains defined by the HFSC battery and the WAIS overlapped to a large extent. In other words, the HFSC tests and the WAIS subtests that make up the composite scores measured many of the same specific cognitive abilities. To elucidate the nature and amount of this overlap, subjects' factor scores computed from the four orthogonally rotated, Session A factors were compared with the WAIS scale scores. Correlations between these sets of composite measures are presented in Table 15. As expected, VIQ correlated most highly with the VERBAL factor, and PIQ was most associated

TABLE 15

Correlations[a] of Session A Factors and WAIS Composite Scores
(N = 448)

Factor	VIQ	PIQ	FIQ
SPATIAL	.26	.56	.45
VERBAL	.62	.33	.58
SPEED	.10	.23	.18
MEMORY	.07	.14	.11
FIRST PCF	.60	.68	.73

a. A correlation greater than .08 is significant at the 5% level, and one of .14 is significant at the .1% level.

with the SPATIAL factor. All four factors were significantly correlated with FIQ.

Since the four factors were orthogonal, it was possible to determine the multiple correlation between the factors and FIQ, PIQ, and VIQ by summing the squared, individual correlations and taking the square root of the summation. This procedure yielded multiple R's of .68 with VIQ, .70 with PIQ, and .76 with FIQ. As a result of these computations, it can be stated that 58% of the variance in FIQ, 49% in PIQ, and 46% in VIQ could be predicted by combining the four Session A factors.

It was hypothesized that the first principal component factor, extracted from the Hawaii tests, provided an estimator of general cognitive ability or Spearman's "g"; and that, for this reason, factor scores computed from the FIRST PCF would correlate well with the Full Scale WAIS measure. Also reported in Table 15, this correlation was .73, indicating that 53% of the variance in FIQ was predictable from FIRST PCF factor scores. The efficacy of prediction of the first component factor was significantly less ($p < .001$) than the combined predictive ability of the four rotated factors, but the latter only accounted for 4.5% more variance.

Although the FIRST PCF was found to be an adequate predictor of FIQ, what remains to be tested is how it compares to the best possible predictor. Since the FIRST PCF only represented 34% of the total variance among the Hawaii tests and the first four components (and rotated factors as well) only 61%, it was feasible that a far better predictor could be obtained by regressing FIQ on the 15 cognitive tests. Results of multiple regressions of FIQ, VIQ, and PIQ are summarized in Table 16. As expected, the multiple correlation coefficients calculated from these analyses were larger than the previously obtained correlations in every case, but not excessively so.

To test the significance of the increased power of prediction when using all 15 cognitive tests, the multiple regression model used above must be restated. Assume that the factor scores obtained from the first principal component had been included in the regression analysis as an additional variable. Further, assume that the 15 cognitive variables in that analysis had been replaced by 15 residual variables, which were created by partialling out the variance due to the first principal component. The resultant multiple correlation coefficient would be identical to that obtained in the

TABLE 16

Summary of Multiple Regressions of WAIS Composite Scores on HFSC Tests

	Dependent Variable		
	VIQ	PIQ	FIQ
Multiple R	.72	.72	.78
Multiple R Squared	.51	.51	.61
Regression Mean Square	219.18	222.39	258.83
Residual Mean Square	208.56	209.70	166.87
$F_{(14,432)}$*	30.27	32.80	44.67

*All three regression equations were significant well beyond the .0001 level.

analysis actually performed; however, one more degree of freedom would have been exhausted because of the addition of a variable, the FIRST PCF factor scores. If the analysis is conceived of in this way, the hypothesis to be tested is that the 15 residual variables make no contribution in predicting the dependent variable, FIQ. That is, 15 hypotheses of the form, $b_i = 0$, are tested concurrently. Statistical significance is determined by computing the F-ratio:

$$F = \frac{(R^2 - r^2)/df1}{(1 - R^2(/df2}$$

where df1 is the number of hypotheses being tested and df2 is N - k - 1 (N being the number of subjects and k being the number of independent variables in the regression equation; see Namboodiri, Carter & Blalock, 1975, pp. 172-175). For testing the relative power of prediction of the four rotated factors, r^2 in the equation is simply replaced by one of the multiple R's already reported, and the degrees of freedom are adjusted accordingly.

The results of this test, when applied to the prediction of VIQ, showed that the regression reported in Table 16 was significantly better than the VERBAL factor by itself ($F_{(15,431)} = 7.94$, $p < .001$) and was better than a regression employing all four cognitive factors ($F_{(15,428)} = 3.32$, $p < .005$). For PIQ, the multiple regression on all 15 tests was far more predictive than the SPATIAL factor ($F_{(15,431)} = 12.13$, $p < .0001$); and it was just significantly better than the multiple regression on four factors would be ($F_{(15,428)} = 1.68$, $p < .05$). Finally, the multiple regression was significantly more predictive of FIQ than was the correlation with the FIRST PCF ($F_{(15,431)} = 5.54$, $p < .001$) and was also better than a regression on all four rotated factors ($F_{(15,428)} = 2.26$, $p < .05$).

Overall, these results statistically confirmed that the best possible predictors of WAIS composite scores, the complete set of HFSC tests, were significantly better than the composite scores derived from the HFSC battery. Another consideration is relevant, however; when dealing with a large sample, statistical significance is not always meaningful. In comparing the amount of variance actually explained by the different prediction equations, it was found that the 15 HFSC variables predicted 6% more variance for VIQ, and only 3% more for PIQ and FIQ, than the four rotated factors.

Likewise, regressing the entire battery on FIQ accounted for only 8% more variance than did the FIRST PCF alone. Thus, the FIRST PCF has been shown to be a significant predictor of the WAIS full scale measure, although a weighted sum of the four factor scores for each subject would provide a 5% better prediction.

A second-order factor measuring general intellectual ability is predicted by the hierarchical model. The measure is derived by factor analyzing the matrix of correlations among obliquely rotated, specific cognitive factors, which, in turn, represent primary mental abilities. An oblique rotation of the four principal components obtained from HFSC tests, and subsequent factor analysis of their intercorrelations, were performed to see whether the second-order, general factor would be a better predictor of FIQ. When factor scores derived from the general measure were correlated with FIQ scores, a value of .70 was obtained. This correlation was no better than that between FIRST PCF and the WAIS criterion, so higher-order factor analyses are not required to obtain an adequate measure of general ability.

All of the measures derived to predict general ability which have been discussed rely on weighted sums of the 15 HFSC test scores. Since such weights would differ slightly from one sample to another, a measure of general ability not dependent on multivariate statistical analyses would be advantageous. Wechsler avoided this problem by equalizing subtest means and variances and using unweighted sums of subtest scores as composite WAIS measures. In a like manner, the sum of standardized HFSC test scores was computed.

A correlation of .72 was obtained between this simple composite measure and FIQ. Thus, although FIRST PCF is a slightly better and more reliable predictor of FIQ, an adequate measure of general ability can be obtained from the HFSC battery without reference to other studies and without the necessity of complicated statistical analyses.

Familial Resemblances for HFSC Tests

The regressions of offspring scores on midparental values was used to measure familial resemblances. Since it is quite possible that environmental correlations between family members' scores are non-zero and positive for some cognitive traits, these regressions provided a measure of the average phenotypic (genetic and/or environmental) similarity across families. Three such estimates of resemblance were computed; $b_{o_i\bar{p}}$, the regression of scores of individual offspring on midparental values (that is, the progeny means were weighted in proportion to the number of progeny measured); $b_{o_1\bar{p}}$, the regression of the oldest offspring's value on midparental scores (losing the information available from additional progeny, but meeting the criteria for a precise statistical test of significance); and $b_{\bar{o}\bar{p}}$, the regression of mid-offspring values on the midparental scores (allowing each family mean the same weight, regardless of the number of progeny tested).

The three estimates of familial resemblance for each of the 15 tests in the HFSC battery and for derived factor scores are presented in Table 17, for the total and AEA samples. For comparison,

TABLE 17

Regressions of Offspring on Midparent for Session A Data and Factors

	Total Sample			AEA Sample			AEA Sample from HFSC (DeFries et al., 1976)
Resemblance Estimate	a±S.E.	b±S.E.	c±S.E.	a±S.E.	b±S.E.	c±S.E.	c±S.E.
Test							
1. VOC	.69±.11	.50±.15	.59±.13	.78±.19	.60±.25	.68±.22	.65±.05
2. VMI	.44±.10	.48±.13	.45±.11	.47±.19	.49±.26	.41±.20	.15±.05
3. TH	.37±.09	.39±.12	.30±.11	.46±.12	.48±.18	.48±.14	.41±.05
4. MR	.47±.10	.45±.13	.37±.11	.24±.16	.21±.19	.23±.18	.42±.05
5. SAM	.55±.10	.63±.14	.50±.12	.59±.15	.75±.17	.57±.17	.38±.04
6. LAD	.19±.09	.24±.11	.16±.10	.31±.14	.31±.21	.31±.16	.22±.05
7. WBE	.31±.10	.22±.14	.31±.12	.22±.14	.15±.19	.24±.16	.40±.04
8. CR	.45±.09	.50±.12	.38±.10	.28±.13	.34±.15	.33±.13	.46±.05
9. VMD	.22±.10	.18±.14	.20±.12	.32±.15	.14±.21	.25±.18	.33±.05
10. PED	.37±.09	.42±.12	.33±.11	.45±.15	.48±.20	.44±.18	.53±.04
11. HP	.47±.10	.56±.12	.37±.11	.60±.17	.59±.17	.56±.17	.43±.05
12. PFB	.44±.08	.56±.11	.45±.10	.44±.13	.49±.18	.46±.16	.53±.04
13. NC	.42±.11	.29±.13	.34±.12	.60±.17	.58±.20	.61±.19	.39±.05
14. SPV	.21±.10	.31±.14	.26±.12	.49±.22	.54±.26	.57±.24	.27±.05
15. PMS	.42±.11	.55±.13	.43±.13	.53±.18	.63±.21	.57±.22	.51±.04
Factors							
Spatial	.45±.08	.62±.11	.46±.10	.40±.13	.46±.15	.42±.14	.57±.04
Verbal	.49±.10	.60±.13	.51±.12	.76±.16	.88±.21	.82±.20	.57±.04
Speed	.55±.10	.54±.14	.46±.13	.66±.16	.88±.19	.67±.19	.41±.05
Memory	.45±.11	.39±.14	.44±.12	.54±.15	.45±.21	.43±.17	.32±.05
First PCF	.38±.09	.45±.11	.38±.10	.58±.14	.61±.18	.64±.17	.61±.04
N	220	118	118	93	51	51	739

a. Three resemblance estimates are reported: (a) the regression of each offspring score on midparent value, effectively weighting by family size; (b) the regression of only the oldest offspring's score on midparental score; and (c) the regression of the average score for offspring within a family on the midparental value, which weights each family the same regardless of the number of offspring.

midchild on midparent regressions ($b_{\bar{o}\bar{p}}$) reported by DeFries and co-workers (1976) are included in the table. The correlation between HFSC and total sample coefficients across the 15 tests was highly significant (Spearman rank correlation = .68, $p < .01$) and that between the HFSC and AEA sets of coefficients was also significant ($r_s = .32$, $p < .05$). There was also no significant difference for any test or factor score when regressions from the two sets of AEA values were compared; however, the resemblance coefficient for VMI was higher ($p < .01$) and that for FIRST PCF was lower ($p < .05$) when total sample regressions were compared to HFSC, AEA values. There is no ready explanation for the difference in VMI regressions. On the other hand, estimated familial resemblance for FIRST PCF was also significantly lower for AJA families than for their AEA counterparts in the HFSC (AJA $b_{\bar{o}\bar{p}} = .43$ compared to .63 for AEAs, when corrected for test reliability). The AJA value is quite comparable to that of the total sample in the present study. Hence, relative to AEAs in the HFSC, the lower resemblance coefficient in the total sample was most probably due to low familial resemblance among the 42% of the sample that was oriental.

For all of the cognitive measures, the three methods of computing familial resemblances gave comparable estimates. That is, no two resemblance coefficients for a given test were significantly different, although two estimates differed at times by as much as 30% (VOC in the total sample). Further examination of the methods uncovered two problems that might occur when comparing different measures of familial resemblance from otherwise similar studies.

First, it can be determined whether a coefficient significantly deviates from a value of zero by dividing the regression of offspring on midparental value by its standard error, which is approximately distributed as Student's t. The different computational methods occasionally led to a contradictory decision regarding this null hypothesis (e.g., LAD, WBE, and VMD in the total sample). Secondly, if an investigator were concerned with the amount of familial resemblance displayed by cognitive tests relative to one another, the use of different methods to estimate resemblance would bias a comparison of studies. Such bias was indicated by calculating a coefficient of concordance over the three sets of 15 resemblance measures reported for the total sample. The test statistic ($W = .58$, $p < .01$) was equivalent to an average Spearman rank correlation coefficient of .78 (Siegal, 1956). An average correlation nearer 1.00 had been expected, since the statistical test was computed across sets of rankings generated from the same data. It appears, however, that the true degree of association between studies may be underestimated when different techniques for computing familial resemblance are used.

Familial Resemblances for the WAIS

In Table 18, the same three methods of ascertaining familial resemblance are presented for WAIS subtests and composite scores in both the total and AEA samples. As was done with the HFSC tests, the concordance across resemblance techniques was computed for the total sample ($W = .94$, $p < .01$) and indicated a significant average

TABLE 18

Regressions of Offspring on Midparent for WAIS Subtests and Composite Scores

Resemblance Estimates[b]	Total Sample			AEA Sample			Canadian Sample[a]
	a±S.E.	b±S.E.	c±S.E.	a±S.E.	b±S.E.	c±S.E.	b±S.E.
Verbal Subtests							
1. Information	.38±.08	.34±.11	.34±.09	.47±.16	.52±.19	.47±.17	.25±.13
2. Comprehension	.49±.09	.41±.12	.44±.10	.27±.17	.35±.23	.31±.18	.26±.13
3. Arithmetic	.42±.10	.48±.13	.33±.10	.48±.14	.60±.18	.49±.16	.45±.12
4. Similarities	.38±.08	.24±.12	.27±.09	.06±.15	-.23±.20	.03±.14	.36±.15
5. Digit Span	.19±.09	.21±.12	.19±.10	.34±.15	.36±.19	.32±.16	.13±.15
6. Vocabulary	.60±.07	.52±.11	.52±.09	.67±.15	.66±.19	.57±.17	.53±.10
Performance Subtests	±						
7. Digit Symbol	.54±.10	.58±.12	.52±.10	.55±.16	.66±.19	.59±.16	.56±.15
8. Picture Completion	.28±.10	.21±.13	.25±.10	.18±.16	-.04±.22	.10±.18	-.12±.20
9. Block Design	.34±.09	.40±.13	.31±.11	.25±.16	.25±.22	.21±.20	.34±.19
10. Picture Arrangement	.06±.08	-.04±.12	-.01±.10	-.05±.13	-.16±.17	-.10±.13	.07±.18
11. Object Assembly	.10±.08	.26±.12	.12±.10	-.08±.10	-.02±.15	-.11±.12	.02±.24
Composite Scales							
VIQ	.50±.08	.46±.12	.41±.09	.51±.15	.50±.20	.49±.17	.55±.09
PIQ	.32±.08	.39±.12	.26±.10	.15±.12	.12±.17	.05±.14	.22±.17
FIQ	.37±.08	.35±.12	.27±.10	.34±.14	.30±.19	.26±.17	.46±.10
N	220	118	118	93	51	51	55

[a] These were reported by Williams (1975).

[b] Three resemblance estimates are reported: (a) the regression of each offspring score on midparent value, effectively weighting by family size; (b) the regression of only the oldest offspring's score on midparental score; and (c) the regression of the average score for offspring within a family on the midparental value, which weights each family the same regardless of the number of offspring.

rank correlation of .91. When testing the null hypothesis, contradictory decisions were indicated among the methods (D, PC and OA in the total sample; D only in the AEA).

Williams (1975) reported the regressions of offspring on midparental scores for WAIS subtests in a Canadian sample. None of his regressions, included in Table 18, differed significantly from comparable values ($b_{o_1\bar{p}}$) computed for the total or AEA samples on any subtest or composite score. In the Canadian sample, resemblances for four subtests (D, PC, PA, and OA) and PIQ were not significantly different from zero, while only three subtests (D, PC, and PA) showed no familial resemblance in the total sample. The Spearman rank correlation between the two sets of regressions was .81, which indicated that the relative magnitudes of resemblance coefficients for subtests were significantly associated ($p < .01$).

Only one other study has provided measures of familial resemblance for WAIS subtests. Block (1968) computed estimates of heritability from intra-class correlations of monozygotic and dizygotic twins. Although these estimates were derived by an entirely different method, the concordance among their rankings and those for both the total and Canadian samples was highly significant ($W = .88$, $p < .01$, $r_{AVE} = .81$). This consistency across studies strongly suggests that the observed differences among subtests were indicative of an underlying, real difference in familial resemblances.

Other Measures of Familial Resemblance

In addition to parent-offspring regressions, it is possible to derive other measures for estimating familial resemblances, which can be used to test additional genetic hypotheses. Six such measures were obtained from data on the WAIS and HFSC batteries. First, the intra-class correlation among offspring was computed as an additional index of familial resemblance. Although based on the expected resemblance among siblings rather than the resemblance of parents and their offspring, twice the value of the sibling correlation yields an estimate of resemblance comparable to a parent-child regression. Another correlation, that between husbands and wives, provided a measure of assortative marriage, which can be roughly defined as the tendency of spouses not to marry at random with regard to the trait being measured. Finally, four parent-child correlations (father-daughter, mother-son, mother-daughter, and father-son) were examined for evidence of sex-linked inheritance. If a variable is sex-linked, the expected pattern of these correlations is: $r_{FD} = r_{MS} > r_{MD} > r_{FS}$. These six measures are presented in Table 19 for the total sample, and in Table 20 for the subsample of AEA families, as computed from HFSC tests and factor scores. Tables 21 and 22 contain the same correlations computed from WAIS data for the total and AEA samples, respectively.

Sex-linkage. Of the 68 sets of parent-child correlations in the four tables, only TH and LAD in the total sample showed the requisite pattern. Confidence intervals for correlation coefficients

TABLE 19

Familial Correlations[a] in the Total Sample: Session A Data

	Father-Son	Mother-Daughter	Mother-Son	Father-Daughter	Husband-Wife	Sibling-Sibling[b]
Tests						
VOC	.28	.33	.34	.22	.03	.32
VMI	.22	.24	.20	.20	-.01	.12
TH	.16	.21	.23	.23	.22	.34
MR	.35	.09	.01	.38	.04	.21
SAM	.31	.11	.29	.36	-.11	.41
LAD	-.01	-.03	.20	.20	.11	.10
WBE	.25	.03	.12	.18	.07	.27
CR	.27	.24	.31	.12	.09	.19
VMD	.18	.10	.03	.19	.11	.19
PED	.33	.26	.41	.19	.36	.28
HP	.12	.28	.34	.24	.03	.16
PFB	.40	.15	.30	.25	.24	.31
NC	.17	.10	.21	.22	-.09	.29
SPV	.01	.25	.15	.06	.10	.20
PMS	.06	.21	.37	.07	.07	.30
Factors						
Spatial	.33	.18	.37	.41	.24	.19
Verbal	.32	.22	.40	.16	.20	.34
Speed	.22	.13	.31	.34	-.03	.43
Memory	.19	.25	.17	.30	.13	.24
First PCF	.26	.11	.37	.30	.19	.30
N	80	81	80	81	118	178[c]

a. Correlations of .22 are significant at the 5% level and correlations of .28 are significant at the 1% level.

b. Intra-class correlations were computed within families of two or more offspring. Values greater than .11 are significant at the .05 level; values greater than .14 at the .01 level.

c. Offspring from 76 families.

TABLE 20

Familial Correlations[a] in the AEA Sample: Session A Data

	Father-Son	Mother-Daughter	Mother-Son	Father-Daughter	Husband-Wife	Sibling-Sibling[b]
Tests						
VOC	.32	.28	.34	.27	.03	.18
VMI	-.15	.46	.32	.00	-.11	.16
TH	.20	.51	.39	.22	.19	.28
MR	.37	.04	-.07	.09	-.02	.03
SAM	.55	.04	.14	.33	-.22	.31
LAD	.28	.08	.03	.29	.07	.10
WBE	.26	-.15	.00	.26	.12	.20
CR	.36	.33	.21	-.03	.23	-.03
VMD	.33	-.07	.01	.46	.21	.16
PED	.24	.31	.39	.21	.32	.19
NP	.21	.42	.12	.39	.09	.15
PFB	.55	.20	.25	.18	.29	.22
NC	.19	.12	.33	.23	-.30	.13
SPV	.08	.22	.20	.14	-.30	.01
PMS	.18	.04	.40	.08	.04	.23
Factors						
Spatial	.51	.24	.20	.29	.27	.14
Verbal	.38	.25	.53	.29	.13	.31
Speed	.31	.15	.34	.35	-.21	.28
Memory	.15	.32	.31	.45	.24	.24
First PCF	.35	.28	.30	.36	.20	.15
N	36	35	36	35	51	78[c]

a. Correlations greater than .33 and .42 are different from zero with $p < .05$ and $p < .01$, respectively.

b. Intra-class correlations greater than .18 and .24 are different from zero at the .05 and .01 levels of significance, respectively.

c. Offspring from 36 families.

TABLE 21

Familial Correlations[a] in the Total Sample: WAIS Scales

	Father-Son	Mother-Daughter	Mother-Son	Father-Daughter	Husband-Wife	Sibling-Sibling[b]
Verbal Tests						
I	.14	.31	.38	.28	.30	.26
C	.16	.35	.21	.21	.03	.18
A	.19	.13	.36	.10	.00	.33
S	.08	.39	.26	.22	.23	.15
D	.36	.01	.24	-.11	.17	.24
V	.34	.38	.51	.26	.25	.38
Performance Tests						
DS	.22	.12	.52	.28	-.06	.16
PC	.03	.18	.17	.10	.07	.09
BD	.26	.18	.24	.06	.07	.20
PA	.05	-.14	.02	.19	.20	.13
OA	.20	-.06	.18	.01	.23	.19
Composite Scores						
VIQ	.27	.24	.46	.16	.18	.36
PIQ	.24	.08	.30	.08	.22	.24
FIQ	.15	.15	.38	.08	.21	.35
N	80	81	80	81	118	178[c]

a. Correlations greater than .22 are significantly different from zero ($p < .05$), and those greater than .28 with $p < .01$.

b. Intra-class correlations were computed within families of two or more offspring. Values greater than .11 are different from zero at the .05 level and those greater than .14 at the .01 level.

c. Offspring from 76 families.

TABLE 22

Familial Correlations[a] in the AEA Sample: WAIS Scales

	Father-Son	Mother-Daughter	Mother-Son	Father-Daughter	Husband-Wife	Sibling-Sibling[b]
Verbal Tests						
I	.03	.24	.57	.37	.12	.29
C	-.03	.15	.14	.09	-.06	.02
A	.43	.07	.45	.20	.02	.41
S	-.42	.28	.15	.15	.23	-.01
D	.44	-.03	.36	.00	.17	.26
V	.20	.08	.72	.24	.17	.32
Performance Tests						
DS	.29	.08	.46	.30	-.07	.03
PC	-.10	.02	.20	.21	.12	.10
BD	.29	.18	.12	.01	.10	.09
PA	-.13	-.21	.00	.17	.27	.04
OA	.15	-.34	.13	-.18	.43	.04
Composite Scores						
VIQ	.12	.08	.57	.27	.12	.36
PIQ	.21	-.01	.13	-.02	.36	-.02
FIQ	.08	-.01	.47	.16	.27	.28
N	36	35	36	35	51	78[c]

a. Correlations greater than .33 and .42 are different from zero with $p < .05$ and $p < .01$, respectively.

b. Intra-class correlations greater than .18 and .24 are significantly different from zero at the .05 and .01 levels, respectively.

c. Offspring from 36 families.

are usually quite large, so it was necessary to ascertain whether the observed patterns for the two tests could be attributed to chance. That the four correlations in each set were not statistically independent was problematical, since testing the homogeneity of coefficients would have violated underlying statistical assumptions. However, Cohen and Cohen (1975) provided a formula for testing the difference between two correlations that share a common variable:

$$t = \frac{(r_{XY} - r_{VY})\ (n - 3)\ (1 + r_{XV})}{[4(1 - r_{XY}^2 - r_{VY}^2 - r_{XV}^2 + 2r_{XY}r_{XV}r_{VY})]^{\frac{1}{2}}}$$

which has n - 3 degrees of freedom.

It was decided to compare the mother and son correlation with that between fathers and sons, and the father-daughter with the mother-daughter correlation. These two tests provided an adequate measure of the presence of sex-linkage (Bock & Kolakowski, 1973). The husband-wife correlation included in the tables was also needed in the test. For LAD in the total sample, r_{MS} was not significantly greater than r_{FS} ($t = 1.48$), nor was r_{FD} different from r_{MD} ($t = 1.35$). The pairs of correlations also were not different for TH in the total sample ($r_{MS} > r_{FS}$, $t = .51$; and $r_{FD} > r_{MD}$, $t = .15$). It can be concluded that there was no indication of sex-linked inheritance for any of the tests or composite scores from the two batteries.

However, one striking trend was evident among the sets of parent-child correlations for all verbal tasks in the total sample except TH. The relationships $r_{MS} > r_{FS}$ and $r_{MD} > r_{FD}$ held consistently for VERBAL, VOC, WBE, PED, and SPV in the HFSC battery and VIQ, I, C, S, and V from the WAIS. This pattern suggests a stronger maternal than paternal influence in the use of language. In the AEA subsample the same trend was found for all HFSC verbal measures; but inconsistent results were observed for WAIS verbal scores.

Spouse correlations. Husband-wife correlations were examined to see whether scores of spouses revealed evidence of assortative marriage for cognitive tasks. For total sample TH, PED, PFB, SPATIAL, VERBAL, I, S, V, PA, OA, and the WAIS composites, spouse scores were significantly related ($p < .05$). Thus some of the tasks appeared relevant in the selection of marital partners, especially for the WAIS, since assortment was indicated by 8 of the 14 measures in that battery. Only PED, PFB, OA, and PIQ exceeded chance amounts of association in the AEA subsample; surprisingly, however, NC and SPV were negatively correlated ($p < .05$). These negative associations would only occur if partners tended to seek mates unlike themselves for the trait. However, the lack of similar coefficients for spouses in the total sample suggests that the finding may be peculiar to the subgroup of AEA couples.

Sibling correlations. Although computed from the same data, different methods for estimating parent-offspring regressions exhibited considerable variation. As measures of familial resemblance, correlations among siblings are quite similar to parent-offspring regressions. In fact, the only major differences involve variance due to genetic dominance and that due to fixed environmental effects. The former is present in sibling correlations, but not in regressions of offspring on midparental values; and the latter probably has a different value among siblings than between parents and their offspring. The two methods were compared to discover how consistently the cognitive measures would be ranked on familial resemblance. Spearman rank correlations were computed between sibling values and all three methods of parent-offspring regression for the total sample. The median correlation was .47 across HFSC tests and .44 across WAIS subtests. None of the six comparisons indicated significantly concordant rankings ($p > .05$). Thus, interpretations of differences within studies must be made with caution. Estimates of resemblance appear to be influenced by even small methodological differences and greatly influenced by large differences.

CHAPTER VI

DISCUSSION

Predictors of General Ability

As hypothesized from factor theory and theories of intelligence, the first principal component factor derived from HFSC tests correlated well with WAIS Full Scale IQ. Smaller associations were also found between the HFSC VERBAL factor and VIQ and between SPATIAL and PIQ.

At first glance, the predictable amount of variation in FIQ (53% by FIRST PCF) seems less substantial than might be expected. However, fairly small decrements were observed when correlations between factor scores and composite WAIS measures were compared to multiple regressions of each composite on the 15 HFSC tests. Since maximal prediction is obtained through multiple regression, relative efficiencies can be determined for the factors with respect to this measure. FIRST PCF was 88% as efficient in predicting FIQ; VERBAL was a 75% efficient predictor of VIQ; and SPATIAL had an efficiency of 61% in predicting PIQ. By combining the four factor scores obtained from a rotated, orthogonal factor analysis of HFSC tests, the relative efficiency in predicting FIQ was increased to 95%; hence, the four factors encapsulated nearly all the predictive power of the 15 HFSC tests.

Another criterion for judging the adequacy of prediction is to compare correlations of FIQ with other accepted measures of intellectual ability to the ones obtained in this study. Such correlations, reported in Table 1, ranged from .45 to .90 with a median of .77. Thus, FIRST PCF (r = .73) and the combined factors (r = .76) are as good at measuring general ability as commonly employed intelligence tests, when this is determined by the correlation with FIQ. Whether general ability is actually a repertoire of acquired, positive attributes acting in concert (Humphreys, 1971) or a unitary, "innate, general, cognitive ability" as proposed by Burt (1955), scores derived from the 15 HFSC tests seem to provide a general measure of cognitive functioning.

However, two aspects of the present study could make any inference to the HFSC, itself, questionable. First, the subsample might not have been representative of the entire HFSC sample; and secondly, measurement properties of the tests might have differed. The latter point was raised because most subjects had been tested with the HFSC battery once or twice before. This led to an overall improvement in scores at retesting; and to complicate matters further, the interval between initial testing and retesting differed considerably among families. Such problems must be addressed before a measure of general ability can be assumed for the HFSC.

A preponderance of oriental families in the subsample was the major difference between it and the total HFSC sample. DeFries et al. (1974) reported "near identity" for AEA and AJA cognitive factor structures in the HFSC, which suggests that the factor structure for Session A data should not have been affected by the relative proportions of oriental and Caucasian families. Factor congruences reported in the 1974 paper ranged from .96 to .99. The lowest congruence computed between AEA and non-AEA factors[1] for this sample was .968 for the SPEED factor, indicating "near identity" for these two groups as well. Since there was no evidence of a group mean difference between AEA and non-AEA subjects for the WAIS Full Scale IQ, predicting this score from standardized factor loadings would be unaffected by the difference in ethnic assortment of the subsample.

Socio-economic status, measured by educational attainment and occupational ratings of fathers, was slightly higher than that in the total HFSC. A positive association between IQ and SES has been well documented (Bajema, 1968; Coleman et al., 1966; Eckland, 1965) so a restriction of sample range could have resulted from testing only higher SES families. This problem would be indicated by smaller standard deviations (SD) than are obtained for unselected samples. Although WAIS subtests suggested that such truncation was present, none of the three composite score SD's differed from those derived for the WAIS standardization sample. Nor did SD's for the 15 HFSC tests reveal any evidence of a restricted range. Finally, since biased sampling would effectively reduce estimates of association between two variables, the presence of range restriction

---

1. These coefficients were computed for this comparison. The five values were: SPATIAL, .992; VERBAL, .991; SPEED, .968; MEMORY, .988; and FIRST PCF, .997.

would only imply that the measures of general ability suggested here would be better predictors in the HFSC than in the subsample. Thus range restriction appears not to be a problem.

Two measurement properties of HFSC tests were evaluated for significant changes to determine the efficacy of the measures at retesting. A high similarity between factor structures showed that the battery was able to define the same independent cognitive functions at Session A as were found by HFSC. Further, correlating subjects' factor scores from Session A with those from initial testing established that discrimination of levels of performance among subjects was also maintained at Session A. It can be concluded, then, that retesting did not significantly affect intercorrelations among tests or relative rankings of scores among subjects.

The only problematical result was an apparent increase in mean scores upon retesting. However, some of this change may be an artifact of the method used for age correction. A number of studies (Schaie & Labouvie-Vief, 1974; Schaie, Labouvie, & Buech, 1973; Schaie & Strother, 1968) have shown that declines with age for cognitive abilities found in cross-sectional studies do not appear in longitudinal studies. More specifically, these authors suggested that differences between age groups cause the apparent declines, since their data showed no decline when subjects were retested as much as 14 years later. In this study, HFSC norms derived from cross-sectional data were used for age adjustment; these correction factors could have differentially inflated scores for subjects, some of whom were retested after nearly two years had elapsed. If present, this bias too would have depressed correlative statistics because of errors dependent on a combination of age and inter-test interval for subjects.

Thus, despite problems related to sampling and measurement, associations between HFSC factors and WAIS scores found in the subsample should also pertain to the HFSC as a whole. Either the first principal component or a weighted sum of factor scores can be used with confidence as an estimate of general intellectual ability by the HFSC.

Comparisons of Familial Resemblances

To explore the effects of procedural differences, three methods of computing regression of offspring on midparental values were employed. One method weighted progeny means by the number of offspring in each family; another utilized only the oldest child; and the last weighted family means equally. Although some ambiguity was found among the methods when testing null hypotheses, the three resemblance estimates did not differ from one another for any WAIS or HFSC variable.

When rankings of resemblance coefficients were compared over methods, concordance coefficients were relatively low. Thus, comparison between studies that used theoretically identical, but computationally different methods, may under-estimate the consistency of variable rankings. Data suggesting similar amounts of variability due to genetic factors for specific cognitive traits must be interpreted cautiously when the conclusion is based on a lack of consistency in the relative magnitudes of resemblance measures. This is especially true when errors inherent in sampling, test reliability, test validity across samples, and procedural differences in testing situations are also considered. Any of these factors could disturb the rank order among resemblance coefficients, while not seriously affecting the statistical validity of the estimates.

Measures of Familial Correlation

Similar caution must be observed when inferences are made from relationships among measures of resemblance within a study. For instance, a peculiar pattern of cross-sex and same-sex, parent-child correlations results from involvement of a major gene located on the X-chromosome. Results of three small studies suggested that spatial visualization is influenced by a sex-linked gene (Stafford, 1961; Hartlage, 1970; Bock & Kolakowski, 1973). Larger, more recent studies (DeFries et al., 1976; Loehlin et al., 1977) have found little evidence supporting this hypothesis. For two variables in the present study, TH and LAD (a test of spatial scanning), patterns of familial correlations were suggestive of sex-linked inheritance. Among 68 sets of correlations, chance alone predicts that two or three patterns of this sort will occur; and in fact, the homogeneity of correlations for both variables was confirmed. Thus, these analyses provide no support for a hypothesis of sex-linkage for spatial ability.

A consistent trend among the parent-child correlations was observed, however. There was greater resemblance between mothers and offspring of either sex than between fathers and offspring of that sex. As used here, the term verbal implies a high (.50 or greater) loading on the HFSC verbal factor or, for WAIS measures, a high loading on the verbal comprehension factor described by Cohen (1957: I, C, S, V, and presumably VIQ). In the total sample all but one of 11 verbal measures showed the pattern. In the smaller AEA sample, three measures were inconsistent.

In general, evidence from twin research suggests that environmental factors act to enhance familial resemblances for verbal tasks. Husén (1963) found that similarity between twins on reading and writing (but not arithmetic) was increased by common learning experience. Nichols (1969) reported that differences in experience decreased twin similarity on the NMSQT, a battery of achievement measures. Verbal tests also had the largest components of between-family variance in analyses of cognitive abilities in twins (Bock & Vandenberg, 1968). This finding was attributed to the sharing of vocabularies by family members, which would lead to a similar knowledge of word usage and similar scores on verbal tests. Finally, Loehlin and Vandenberg (1968) examined genetic and environmental covariation in a battery of cognitive measures. They reported a second-order verbal factor in the environmental covariation and a general factor in the genetic. Other research (Smith, 1965; Scarr, 1969) indicated that such environmental influence may affect monozygotic more than dizygotic twins. The present data further suggest that mothers are a larger source of influence than fathers.

Since parent-child correlations were reported for HFSC tests by DeFries et al. (1976), their results were examined for evidence of a similar trend. Among AEA families the pattern was consistent for every verbal measure with the surprising exception of the verbal factor score. On it, the correlation between mothers and daughters equaled that between fathers and daughters. The trend was not apparent, however, in AJA correlations from the same report. Data collected with HFSC tests in Korea also revealed no consistency among correlations for verbal scores (Park, 1975). And, as a pessimist might predict, nearly the opposite tendency was observed for parent-son correlations from WAIS data collected in Canada (Williams, 1975).

Evidence for a greater maternal, environmental effect on verbal tasks might be inferred from the present data and that of DeFries and his co-workers. However, contradictory evidence from other parent-child data definitely limits the supposition; only Americans of European ancestry seem to exhibit higher mother-child correlations for verbal tasks.

Familial Resemblance for IQ Measures

Estimated familial resemblances for WAIS Full Scale IQ were comparable to the value of .46 reported by Williams (1975). Since the regression of offspring on midparental value only provides an upper-limit estimate of heritability in its "narrow sense," these results suggest that the true heritability of intelligence may be lower than reported, "broad heritability" values of .80 (Jensen, 1969) and .71 (Jinks & Fulker, 1970) and nearer the estimate of .45 presented by Jencks (1973). This conclusion is supported by estimated resemblances for general ability obtained from the HFSC battery, which were comparable to values reported by DeFries and his colleagues (1976).

Most estimates of heritability for intelligence have been derived from comparisons between identical and fraternal twins. Unfortunately, a number of factors can seriously bias such estimates. For instance, it has been suggested that identical twins may be more similar than fraternal because they are treated more alike by individuals in their environment. Although there is little evidence of an effect of this type for cognitive tests (Vandenberg, 1976; Loehlin & Nichols, 1976), it would cause estimates of heritability to be inflated, if present. Genetic dominance and epistasis are other factors that lead to overestimation of "narrow sense" heritability in twin studies. Kempthorne and Osborne (1961) identified 11 such factors that can affect MZ twin correlations and cautioned against estimating genetic variance from twin data alone.

Estimates of heritability obtained by regressing offspring scores on midparental values are also biased by a variety of factors, but there are fewer drawbacks for this method than for twin comparisons. Three of the major sources of variation that can affect parent-child resemblances for cognitive measures are: environmental covariation between parents and offspring; differential validity or reliability for measures across generations; and epistatic variance due to interactions among genes that influence the trait. The first of these biases is difficult to assess and must be assumed to vary from measure to measure. The second is the easiest to ascertain of the three and, if necessary, different correction factors can be applied to generational data before proceeding with the computation of resemblance coefficients. Lastly, epistatic effects are not easily estimated from human data; but with respect to the comparison being made, epistasis will bias estimates of heritability derived from twin comparisons to a much greater degree than estimates obtained from parent-child regressions. Overall, regressions of offspring on parental values more accurately estimate components of additive genetic variance than do methods of twin comparison. This advantage strengthens the conclusion that the heritable component in IQ is probably lower than previously estimated from comparisons of MZ and DZ twins.

Familial Resemblances for Specific Abilities

Seemingly, in support of the thesis of roughly equal degrees of genetic determination for specific abilities, no significant differences were found between any two WAIS or HFSC parent-child regressions. Lack of significance may have been due, in part, to the comparatively small sample of the present study, since trends in the data were fairly consistent with differences in regressions found by DeFries and his co-workers (1976) for HFSC tests. However, all three tests showing greater familial resemblance in the larger study correlated highly with the first principal component. They were not measures of specific abilities as much as they were measures of general cognitive ability. It is only safe to assume that additive environmental effects may inflate estimates of familial resemblance for measures of general ability, and that the DeFries results may be biased to some extent by these factors.

Although no pair-wise differences were found among resemblance coefficients, striking evidence was in support of the contention that specific cognitive abilities are determined to different degrees by genetic factors. Relative rankings of resemblance coefficients showed high concordance with results of previous researchers. In fact, consistency among studies was almost as great as the concordance of various computational methods applied to the data of the present study. Only remarkably consistent environmental determinants, specific to each cognitive ability, provide an adequate counter-explanation. No evidence has yet been presented supporting such an alternative.

The present study does provide some evidence consistent with the hypothesis that specific cognitive abilities are differentially influenced by hereditary factors. However, this support was not definitive because statistical differences in familial resemblance among cognitive abilities could not be demonstrated.

Suggestions for Further Research

It should be obvious from the previous discussion that large family studies are necessary if the question of varying genetic determination for cognitive traits is to be resolved. Ideally, these studies would include parents, twins, and non-twin siblings so that analyses of the type proposed by Jinks and Fulker (1970) and Elston and Gottesman (1968) could be applied. Such analyses, which concurrently use parent-offspring, sibling-sibling, and MZ versus DZ resemblances to maximize the accuracy of parameter estimation in genetic models, are both appealing and powerful. Results could be further enhanced by combining these techniques with an adoption paradigm. The use of both biological and adoptive families provides accurate determination of systematic environmental effects as well as the extent of gene-environment interaction and correlation (Plomin, DeFries, & Loehlin, 1977). By adding estimation of these factors to the precise estimates of genetic components obtained through analyses of multiple relationships, the question of whether or not specific cognitive abilities are similarly influenced by heredity would be answered.

Unfortunately, a study of the type proposed requires extensive planning and considerable effort at great expense. Until the definitive study has been designed, funded, and executed, it may be possible to obtain a partial answer by investigating genetic cross-correlations (Reeve, 1955; Falconer, 1960) among cognitive traits. This method could be applied to extant data to establish whether some abilities have similarly moderate, but genetically unique determination. In other words, some tests may show little or no common genetic variance, although they have similar familial resemblances and are correlated. Thus, the method would provide evidence of either varying environmental factors or different clusters of genes determining familial resemblances of cognitive abilities.

Such evidence would do more to elucidate those heritable cognitive factors which contribute to general cognitive ability than prove or disprove if the factors themselves differ in heritability. Toward this goal, results of the present study have established that a measure of general ability can be extracted from the Hawaii Family Study of Cognition battery that will adequately discriminate levels of ability in the entire HFSC sample. With this measure, the search for determinants and correlates of general intellectual ability in the HFSC can continue with confidence.

REFERENCES

Adams, B., Ghodsian, M., & Richardson, K. Evidence for a low upper limit of heritability of mental test performance in a national sample of twins. Nature, 1976, 263, 314-316.

Bajema, C. J. A note on the interrelations among intellectual ability, educational attainment, and occupational achievement: A follow-up study of a male Kalamazoo public school population. Sociology of Education, 1968, 41, 317-319.

Binet, A., & Simon, T. H. Method of measuring the development of the intelligence of young children. Chicago: Chicago Medical Book Company, 1915.

Block, J. B. Hereditary components in the performance of twins on the WAIS. In S. G. Vandenberg (Ed.), Progress in human behavior genetics, Recent reports on genetic syndromes, twin studies and statistical advances. Baltimore: Johns Hopkins Press, 1968. Pp. 221-228.

Bock, R. D., & Kolakowski, D. Further evidence of sex-linked major-gene influence on human spatial visualizing ability. American Journal of Human Genetics, 1973, 25, 1-14.

Bock, R. D., & Vandenberg, S. G. Components of heritable variation in mental test scores. In S. G. Vandenberg (Ed.), Progress in human behavior genetics. Baltimore: Johns Hopkins Press, 1968. Pp. 233-260.

Boring, E. G. Intelligence as the tests test it. New Republic. 1923, 35, 35-37.

Brazziel, W. F. A letter from the south. In Environment, heredity, and intelligence, Harvard Educational Review, Cambridge, Reprint 2, 1969.

Buros, O. K. The sixth mental measurements yearbook. Highland Park, N.J.: Gryphon Press, 1965.

Burt, C. L. Reports of consultative committee on psychological tests of educable capacity. London: H. M. Stationery Office, 1924.

Burt, C. L. The evidence for the concept of intelligence. British Journal of Educational Psychology, 1955, 25, 158-177.

Cattell, R. B. The structure of intelligence in relation to the nature-nurture controversy. In R. Cancro (Ed.), Intelligence, genetic and environmental influences. New York: Grune and Stratton, 1971.

Cohen, J. A comparative factor analysis of WAIS performance for four age groups between 18 and 80. American Psychologist, 1956, 11, 449.

Cohen, J. The factorial structure of the WAIS between early adulthood and old age. Journal of Consulting Psychology, 1957, 21, 283-290.

Cohen, J., & Cohen, P. Applied multiple regression/correlation analysis for the behavioral sciences. Hillsdale, N.J.: Lawrence Erlbaum Associates, Publishers, 1975.

Cole, D., & Weleba, L. Comparison data on the Wechsler Bellevue and the WAIS. Journal of Clinical Psychology, 1956, 12, 198-199.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Wainfeld, F. D., & York, R. L. Equality of educational opportunity. Washington, D.C.: U. S. Government Printing Office, 1966.

Conry, R., & Plant, W. T. WAIS and group test prediction of an academic success criterion: High school and college. Educational and Psychological Measurement, 1965, 25, 493-500.

Coons, W. H., & Peacock, E. P. Inter-examiner reliability of the WAIS with mental hospital patients. Ontario Psychological Association Quarterly, 1959, 12, 33-37.

Cronbach, L. J. Coefficient alpha and the internal structure of tests. Psychometrica, 1951, 16, 297-334.

Davies, A. D. M., & Davies, M. G. The difficulty and graded scoring of Elithorn's perceptual maze test. British Journal of Psychology, 1965, 56, 295-302.

DeFries, J. C., Ashton, G. C., Johnson, R. C., Kuse, A. R., McClearn, G. E., Mi, M. P., Rashad, M. N., Vandenberg, S. G., & Wilson, J. R. Parent-offspring resemblance for specific cognitive abilities in two ethnic groups. Nature, 1976, 261, 131-133.

DeFries, J. C., Vandenberg, S. G., & McClearn, G. E. Genetics of specific cognitive abilities. Annual Review of Genetics, 1976, 10, 179-207.

DeFries, J. C., Vandenberg, S. G., McClearn, G. E., Kuse, A. R., Wilson, J. R., Ashton, G. C., & Johnson, R. C. Near identity of cognitive structure in two ethnic groups. Science, 1974, 183, 338-339.

Eaves, L. J. Assortative mating and intelligence: An analysis of pedigree data. Heredity, 1973, 30, 199-210.

Eber, H. W. Measurement instruments. In R. M. Dreger (Ed.), Multivariate personality research: Contributions to the understanding of personality in honor of Raymond B. Cattell. Baton Rouge, La.: Claitor's Publishing Division, 1972. Pp. 51-71.

Eckland, B. K. Academic ability, higher education, and occupational mobility. American Sociological Review, 1965, 30, 735-746.

Elithorn, A., Kerr, M., & Mott, J. A group version of a perceptual maze test. British Journal of Psychology, 1960, 51, 19-26.

Elston, R. C., & Gottesman, I. I. The analysis of quantitative inheritance simultaneously from twin and family data. American Journal of Human Genetics, 1968, 20, 512-532.

Erlenmeyer-Kimling, L., & Jarvik, L. F. Genetics and intelligence: A review. Science, 1963, 142, 1477-1479.

Falconer, D. S. Introduction to quantitative genetics. New York: Ronald Press, 1960.

Feldman, M. W., & Lewontin, R. C. The heritability hang-up. Science, 1975, 190, 1163-1168.

French, J. W., Ekstrom, R. G., & Price, L. A. Manual for a kit of reference tests for cognitive factors. Princeton, N.J.: Educational Testing Service, 1963.

Galton, F. Hereditary genius: An inquiry into its laws and consequences. London: Macmillan, 1869.

Giannell, A. S., & Freeburne, C. M. The comparative validity of the Wechsler and Stanford Binet with college freshmen. Educational and Psychological Measurement, 1963, 23, 557-567.

Granick, S. Brief tests and their interrelationships as intellectual measures of aged subjects. American Psychological Association Proceedings, 1971, 79, 599-600.

Guilford, J. P. The structure of intellect. Psychological Bulletin, 1956, 53, 267-293.

Guilford, J. P. Three faces of intellect. American Psychologist, 1959, 14, 469-479.

Hartlage, L. C. Sex-linked inheritance of spatial ability. Perceptual and Motor Skills, 1970, 31, 610.

Hawaii Family Study of Cognition Staff Report, Report to co-investigators of the cognition grant for meeting held June 20, 1976, unpublished.

Heim, A. W. The appraisal of intelligence. London: Methuen, 1954. Ch. 12.

Heim, A. W. Intelligence and personality: Their assessment and relationship. Baltimore: Penguin Books, 1970.

Holzinger, K. J. The relative effect of nature and nurture influences on twin differences. Journal of Educational Psychology, 1929, 20, 241-248.

Horn, J. L. The structure of intellect: Primary abilities. In R. M. Dreger (Ed.), Multivariate personality research: Contributions to the understanding of personality in honor of Raymond B. Cattell. Baton Rouge, La.: Claitor's Publishing Division, 1972. Pp. 451-511.

Horn, J. L. Human abilities: A review of research and theory in the early 1970's. Unpublished manuscript, 1976.

Humphreys, L. G. Theory of intelligence. In R. Cancro (Ed.), Intelligence, genetic and environmental influences. New York: Grune and Stratton, 1971. Pp. 31-42.

Husén, T. Intra-pair similarities in the school achievements of twins. Scandinavian Journal of Psychology, 1963, 4, 108-114.

Jencks, C. Inequality: A reassessment of the effect of family and schooling in America. New York: Basic Books, 1972.

Jensen, A. R. How much can we boost IQ and scholastic achievement? In Environment, heredity and intelligence, Harvard Educational Review, Reprint 2, Cambridge, 1969.

Jensen, A. R. Genetics and education. New York: Harper & Row, 1972.

Jensen, A. R. Educability and group differences. New York: Harper & Row, 1973.

Jinks, J. L., & Eaves, L. J. IQ and inequality. Nature, 1974, 248, 287-289.

Jinks, J. L., & Fulker, D. W. Comparison of the biometrical genetical, MAVA, and classical approaches to the analysis of human behavior. Psychological Bulletin, 1970, 73, 311-349.

Johnson, R. C., DeFries, J. C., Wilson, J. R., McClearn, G. E., Vandenberg, S. G., Ashton, G. C., Mi, M. P., & Rashad, M. N. Assortative marriage for specific cognitive abilities in two ethnic groups. Human Biology, 1976, 48, 343-352.

Jones, L. V. A factor analysis of the Stanford-Binet at four age levels. Psychometrika, 1949, 14, 299-331.

Kaiser, H. F. The application of electronic computers to factor analysis. Educational and Psychological Measurement, 1960, 20, 141.

Kamin, L. J. The science and politics of I.Q. Potomac, Md.: Lawrence Erlbaum Associates, 1974.

Kangas, J., & Bradway, K. Intelligence at middle age. Developmental Psychology, 1971, 5, 333-337.

Kelley, T. L. Crossroads in the mind of man: A study of differential mental abilities. Palo Alto, Calif.: Stanford University Press, 1928.

Kempthorne, O., & Osborne, R. H. The interpretation of twin data. American Journal of Human Genetics, 1961, 13, 320-339.

Kennedy, W. A., Willcutt, H., & Smith, A. Wechsler profiles of mathematically gifted adolescents. Psychological Reports, 1963, 12, 259-262.

Layzer, D. Heritability analyses of IQ scores: Science or numerology? Science, 1974, 183, 1259-1266.

Lewontin, R. C. Genetic aspects of intelligence. Annual Review of Genetics, 1975, 9, 387-405.

Loehlin, J. C., & Nichols, R. C. Heredity, environment and personality: A study of 850 sets of twins. Austin: University of Texas Press, 1976.

Loehlin, J. C., Sharan, S., & Jacoby, R. In pursuit of the "spatial gene": A family study. In press, 1977.

Loehlin, J. C., & Vandenberg, S. G. Genetic and environmental components in the covariation of cognitive abilities: An additive model. In S. G. Vandenberg (Ed.), Progress in human behavioral genetics. Baltimore: Johns Hopkins Press, 1968. Pp. 261-285.

Lord, F. M., & Novick, M. R. Statistical theories of mental test scores. Reading, Mass.: Addison-Wesley Publishing Company, 1968.

Lubin, B., Wallis, R. R., & Paine, C. Patterns of psychological test usage in the United States: 1935-1969. Professional Psychology, 1971, 2, 70-74.

McClelland, D. Testing for competence rather than for intelligence. American Psychologist, 1973, 61, 1-14.

Miles, T. R. Contributions to intelligence testing and the theory of intelligence. I. On defining intelligence. British Journal of Educational Psychology, 1957, 27, 153-165.

Namboodiri, N. K., Carter, L. F., & Blalock, H. M. Jr. Applied multivariate analysis and experimental designs. New York: McGraw Hill Inc., 1975.

Neuringer, C. The form equivalence between the WB IQ scale, Form I, and the WAIS. Educational and Psychological Measurement, 1963, 23, 755-764.

Nichols, R. C. The inheritance of general and specific ability. In M. Manosevitz, G. Lindzey, & D. D. Thiessen (Eds.), Behavioral genetics: Method and research. New York: Appleton-Century-Crofts, 1969. Pp. 318-327.

Olsen, I. A., & Jordheim, G. D. Use of WAIS in a student counseling center. Personnel and Guidance Journal, 1964, 42, 500-507.

Park, J. Y. A study of multivariate cognition in Korea in relation to environmental and hereditary influence. Unpublished doctoral dissertation, University of Hawaii, 1975.

Paterson, D. G., Elliot, R. M., Anderson, L. D., Toops, H. A., & Heidbreder, E. Minnesota Paper Form Board Test. New York: Psychological Corp., 1930.

Plant, W. T., & Lynd, C. A validity study and a college freshman norm group for the WAIS. Personnel and Guidance Journal, 1959, 37, 578-580.

Plomin, R., DeFries, J. C., & Loehlin, J. C. Genotype-environment interaction and correlation in the analysis of human behavior. Psychological Bulletin, 1977, in press.

Quershi, M. Y. The optimum limits of testing on the Wechsler intelligence scales. Genetic Psychology Monograph, 1968, 78, 141-190.

Quershi, M. Y., & Miller, J. M. The comparability of the WAIS, WISC and WB II. Journal of Educational Measurement, 1970, 7, 105-111.

Rao, D. C., Morton, N. E., & Yee, S. Analysis of family resemblance II, a linear model of familial correlation. American Journal of Human Genetics, 1974, 26, 331-359.

Raven, J. C. Standardization of progressive matrices, 1938. British Journal of Medical Psychology, 1941, 19, 137-150.

Reeve, E. C. R. The variance of the genetic correlation coefficient. Biometrics, 1955, 11, 357-374.

Reiss, A. J. Occupations and social status. Glencoe, Ill.: Free Press, 1961.

Richardson, K., & Spears, D. Race and intelligence: The fallacies behind the race-IQ controversy. Baltimore: Penguin Books, 1972.

Rimoldi, H. J. A. The central intellective factor. Psychometrica, 1951, 16, 75-102.

Scarr, S. Environmental bias in twin studies. In M. Manosevitz, G. Lindzey, & D. D. Thiessen (Eds.), Behavioral genetics: Method and research. New York: Appleton-Century-Crofts, 1969.

Schaie, K. W., Labouvie, G. V., & Buech, B. U. Generational and cohort-specific differences in adult cognitive functioning: A fourteen year study of independent samples. Developmental Psychology, 1973, 9, 151-166.

Schaie, K. W., & Labouvie-Vief, G. V. Generational versus ontogenetic components of change in adult cognitive behaviors: A fourteen-year cross-sequential study. Developmental Psychology, 1974, 10, 305-320.

Schaie, K. W., & Strother, C. R. The effects of time and cohort differences on the interpretation of age changes in cognitive behavior. Multivariate Behavioral Research, 1968, 3, 259-294.

Schwartz, M., & Schwartz, J. Evidence against a genetical component to performance in IQ tests. Nature, 1974, 248, 84-85.

Shaw, D. J. Sexual bias in the WAIS. Journal of Consulting Psychology, 1965, 29, 590-591.

Shaw, D. J. Factor analysis of the collegiate WAIS. Journal of Consulting Psychology, 1967, 31, 217.

Shepard, R. N., & Metzler, J. Mental rotation of three-dimensional objects. Science, 1971, 171, 701-703.

Siegal, S. Nonparametric statistics for the behavioral sciences. New York: McGraw Hill, 1956.

Smith, R. T. A comparison of socio-environmental factors in monozygotic and dizygotic twins, testing an assumption. In S. G. Vandenberg (Ed.), Methods and goals in behavior genetics. New York: Academic Press, 1965. Pp. 45-61.

Spearman, C. General intelligence, objectively determined and measured. American Journal of Psychology, 1904, 15, 201-293.

Spearman, C. The abilities of man. New York: Macmillan, 1927.

Stafford, R. E. Sex differences in spatial visualization as evidence of sex-linked inheritance. Perceptual and Motor Skills, 1961, 13, 428.

Tallman, G. G. A comparative study of identical and non-identical twins with respect to intelligence resemblances. In 27th Yearbook of the National Society on the Study of Education (Part I). Bloomington, Ill.: Public School Publishing, 1928. Pp. 83-86.

Terman, L. M., Thorndike, E. L., Thurstone, L. L., et al. Intelligence and its measurement: A symposium. Journal of Educational Psychology, 1921, 12, 123-154, 195-216.

Thomson, G. The factorial analysis of human ability. London: University of London Press, Ltd., 1950.

Thurstone, L. L. Primary mental abilities. Chicago: University of Chicago Press, 1938.

Thurstone, L. L., & Thurstone, T. G. Factorial studies of intelligence. Chicago: University of Chicago Press, 1941.

Thurstone, L. L., & Thurstone, T. G. S.R.A. Primary mental abilities. Chicago: Science Research Associates, 1947.

Vandenberg, S. G. The nature and nurture of intelligence. In D. C. Glass (Ed.), Genetics. New York: Rockefeller University Press, 1968. Pp. 3-58.

Vandenberg, S. G. A twin study of spatial ability. Multivariate Behavioral Research, 1969, 4, 273-294.

Vandenberg, S. G. Hereditary factors in psychological variables in man, with a special emphasis on cognition. In J. N. Spuhler (Ed.), Genetic diversity and human behavior. Chicago: Aldine, 1974. Pp. 99-133.

Vandenberg, S. G. Twin studies. In A. R. Kaplan (Ed.), *Human behavior genetics*. Springfield: Thomas Books, 1976. Pp. 90-150.

Vandenberg, S. G., & Kuse, A. R. Mental rotations, a test of three-dimensional spatial visualization. Submitted.

Vandenberg, S. G., Meredith, W., & Kuse, A. R. Two higher order factors in intelligence test scores. A partial replication of the Cattell-Horn theory of fluid and crystallized intelligence. Unpublished manuscript.

Wechsler, D. *Measurement of adult intelligence*. 3rd ed. Baltimore: Williams and Wilkins, 1944.

Wechsler, D. *Manual for the Wechsler adult intelligence scale*. New York: The Psychology Corporation, 1955.

Wechsler, D. *The measurement and appraisal of adult intelligence*. 4th ed. Baltimore: Williams and Wilkins, 1958.

Whiteman, M. The performance of schizophrenics in social concepts. *Journal of Abnormal and Social Psychology*, 1954, *49*, 266-271.

Williams, T. Family resemblance in abilities: The Wechsler scales. *Behavior Genetics*, 1975, *5*, 405-409. (a)

Williams, T. The Wechsler scales: Parents and (male) children. *Journal of Educational Measurement*, 1975, *12*, 119-128. (b)

Wilson, J. R., DeFries, J. C., McClearn, G. E., Vandenberg, S. G., Johnson, R. C., & Rashad, M. N. Cognitive abilities: Use of family data as a control to assess sex and age differences in two ethnic groups. *International Journal on Aging and Human Development*, 1975, *6*, 261-276.

Zimmerman, I. L., & Woo-Sam, J. M. *Clinical interpretation of the Wechsler adult intelligence scale*. New York: Grune and Stratton, 1973.

APPENDIXES

APPENDIX A

DESCRIPTIONS OF HFSC BATTERY TESTS

The Hawaii Family Study of Cognitive Test Battery is composed of 15 cognitive measures. These were selected from a battery of 46 measures administered to a pilot sample of 172 individuals (Vandenberg, Meredith, & Kuse, unpublished). These are described in the order of administration. Sources, scoring procedures, and any modifications of the measures are mentioned.

1. Vocabulary Test. (PMA battery; copyright Science Research Associates, Inc.)

The Vocabulary test consisted of fifty multiple choice items, on which subjects were allowed to work for three minutes. Each item consisted of a stimulus word at the left of a line with four possible synonyms to the right. Subjects were instructed to circle the correct synonym. The score computed was the number of correct responses adjusted for guessing.[1]

2. Visual Memory Test (Immediate Recall). (Constructed by BBL staff using illustrations from the Golden Picture Dictionary.)

Subjects were instructed to memorize two pages of illustrations of common objects. Each page contained 20 illustrations.

---

[1] The standard correction for guessing, $S = R - \frac{W}{N - 1}$, was employed for many of the tests. S is the corrected score; R is the number of items correctly answered; W, the number of items answered incorrectly; and N, the number of responses possible for each item.

After one minute of study, subjects were to turn the page and in one minute circle those illustrations recalled from the study period. The two test pages also contained 20 illustrations each; however, only 10 illustrations per page had been presented as stimulus items. The other 10 drawings on each page were novel distraction items. The score derived was the number of illustrations correctly circled, less the number of distractors circled.

3. Things Categories Test. (ETS Kit of Reference Tests for Cognitive Factors, revised 1963; copyright Educational Testing Service.)

The test was administered in two parts of three minutes each. Subjects were instructed to list the names of all "things that are often round" in Part 1, and to list "things that are often metal" in Part 2. The second part of the test was a modification of the ETS procedure, which requires subjects to name objects that are "blue." Thus, both the content of the second part and the wording on both parts were altered. The scoring procedure also differed from that suggested by ETS. Rather than simply counting the number of names of things listed, an effort was made to discount inappropriate responses. For example, "house" was not counted as a correct response on either part of the test. A general criterion for scoring responses was that the object named be round, or metal, at least half the time. The number of acceptable responses on both parts was the score.

4. Mental Rotations Test. (Adapted from the Shepard and Metzler (1971) figures; Vandenberg and Kuse, submitted.)

The test is composed of figures constructed from ten blocks. A stimulus figure was presented at the left of a line, and four possible congruent figures are shown to the right. Two of the four response choices represent the stimulus as it would appear if it had been rotated to a new orientation. The other two choices are either a mirror-image representation of the stimulus item or an entirely different arrangement of blocks. Twenty stimuli are presented, so 40 correct responses are possible. The score for this test was the number of correct responses in a ten minute period, minus the number of incorrect responses. Negative scores were allowed.

5. Subtraction and Multiplication Test. (ETS Kit of Reference Tests for Cognitive Factors, revised 1963; copyright Educational Testing Service.)

The test consists of two parts, each containing 60 items. The format of the test alternates rows of 10 items by subtracting two-digit numbers from two-digit numbers, with 10 items of multiplying two-digit numbers by single-digit numbers. The score was the sum of the number of correct responses on both parts.

6. Lines and Dots. (Selected mazes from Elithorn et al., 1960.)

Ten of the Elithorn Perceptual Mazes were chosen for inclusion in the battery. Each maze consists of a V-shaped grid of dotten lines. Heavy black dots are scattered on the grid at some of the intersections of the dotted lines. Subjects were instructed to connect a specified number of heavy black dots according to three rules: (1) start at the bottom of the maze; (2) always follow the dotted lines through the maze; and (3) always move toward the top of the maze without retracing steps. A method allowing partial credit, and based on the difficulty of each maze, was used for scoring (Davies & Davies, 1965). Although Davies and Davies suggest that the method may be inappropriate when subjects know the number of heavy dots to be connected, the method yielded scores correlating at least as highly, and often more highly, with other spatial tests than did alternative scoring procedures.

7. Word Beginnings and Endings Test. (ETS Kit of Reference Tests for Cognitive Factors, revised 1963; copyright Educational Testing Service.)

The test consists of two three-minute parts. On each part the subject is required to write as many words as possible that begin with one given letter and end with another. The score used was the number of words written that also appeared in a large unabridged dictionary. Obvious misspellings were accepted, but nonsensical words were not scored as being correct.

8. Card Rotations Test. (ETS Kit of Reference Tests for Cognitive Factors, revised 1963; copyright Educational Testing Service.)

The test consists of two parts of 14 items with three minutes allowed for each part. Every item contains an irregularly shaped stimulus to the left and, to its right, six other drawings of the same item rotated in the plane of the paper. Some of the rotated drawings are identical to the criterion, and the rest are mirror-image representations of the criterion. The subject must decide which drawings are merely rotated and which are rotated mirror-images. The score used was the number of correctly identified drawings, corrected for guessing.

9. Visual Memory Test (Delayed Recall). (Constructed by BBL staff using illustrations from the Golden Picture Dictionary.)

The remaining 20 illustrations memorized for the Visual Memory Test-Immediate Recall (test 2 in the battery) were presented with 20 new distractors. The subjects were instructed to circle those illustrations which they had memorized earlier. One minute was sufficient to complete the task. The score computed was the number of illustrations correctly identified, minus the number of distractors circled.

10. Pedigrees Test. (PMA battery; copyright Science Research Associates, Inc.)

A pedigree of a family is presented at the top of the test page. Below are a set of 20 questions with five possible correct answers referring to the pedigree. The test consists of two pedigrees and sets of questions to be completed in four minutes. As a score, the number of questions correctly answered, corrected for guessing, was used.

11. Hidden Patterns Test. (ETS Kit of Reference Tests for Cognitive Factors, revised 1963; copyright Educational Testing Service.)

Each item consists of a configuration of straight line segments, some of which contain a specified embedded pattern. The task is to circle those configurations in which the embedded pattern occurs. Two minutes are permitted for each of two parts containing 200 configurations. The number of configurations correctly identified, minus those patterns circled which did not contain the specified configuration, was used as the score for the test.

12. Paper Form Board Test. (Adapted from the Minnesota Paper Form Board Test; Paterson, Elliot, Anderson, Toops, and Heidbreder, 1930.)

For each item in this test, an outline of a simple geometric figure is presented. To its right are two to four pieces that would result if the simple form had been cut up. The subject is instructed to show how the figure should be cut to form the pieces by drawing lines on it. The score is the number of items entirely correct. Subjects were allowed three minutes for 28 items.

13. Number Comparisons Test. (ETS Kit of Reference Tests for Cognitive Factors, revised 1963; copyright Educational Testing Service.)

Pairs of multi-digit numbers are presented to the subject, who must decide whether the two numbers of each pair are the same or different. The subject is instructed to mark the pairs in which the numbers are not the same. One and one-half minutes were allowed for the 48 items in each of the two parts of the test. The score is the number of pairs correctly identified, less the number of pairs incorrectly identified.

14. Whiteman Test of Social Perception. (Adapted from Whiteman, 1954.)

Each of the 20 items in the test consisted of a set of four sketches. The task for subjects consisted of two parts: first, to identify which sketch did not illustrate the same social concept as the other three; and secondly, to describe with a short phrase the social concept depicted in the remaining three drawings. The score used for the test was the number of social concepts correctly described in the ten minute time period.

15. Progressive Matrices Test (shortened form). (Selected with permission from the Progressive Matrices Test of Raven, 1941.)

The test consists of 29 items selected from Raven's Progressive Matrices. In the order of presentation these are items: A5, A8, A12, A11, B5, B6, B8, B10, B12, C3, C6, C8, C11, C12, D2, D3, D6, D8, D10, E1, E2, E5, E6, E10, E11, E7, E8, E9, and E12. During the first year of the HFSC item A1 was used as an example. After the first year, A1 was replaced with item B4, since it was found to be more representative of the test in its shortened form. In a pilot sample of 172 individuals, the score derived from the shortened version was found to correlate .89 with the score from the entire 60 items. This value is comparable to the reported reliability for the entire set of Progressive Matrices. A second pilot study of 46 individuals confirmed that item efficacy and correlations with other measures were unaffected by using the shortened version. The number of items correctly answered in 20 minutes was used as the score.

APPENDIX B

DESCRIPTIONS OF WAIS SUBTESTS

Only brief descriptions of WAIS subtests are presented; more detailed information on items, scoring procedures, and interpretation of scores has been prepared by Wechsler (1955, 1958) and Zimmerman and Woo-Sam (1973). In general, the verbal subtests are presented orally, and subjects must also respond orally. Performance subtests require combinations of oral response and manipulation of test materials by the administrator, the subject, or both.

Verbal Subtests

Information. The first subtest is composed of 29 items of general information that can be obtained experientially. An attempt was made to avoid the use of specialized or academic knowledge so that the test would be applicable to the entire adult population.

Comprehension. The 14 Comprehension items were designed to measure practical judgement and common sense. Subjects are required to explain the meaning of proverbs, what should be done under a given set of circumstances, and the reason for common practices (such as the necessity of a marriage license).

Arithmetic. Fourteen problems involving addition, subtraction, multiplication, division, and fractions must be solved without the use of paper and pencil. These items are similar to arithmetic problems encountered in elementary school.

Similarities. This measure assesses the extent to which subjects have assimilated similarities and differences between objects, facts, or concepts. The subject must identify an essential way that two things are alike.

Digit Span. On this measure of immediate auditory recall, series of from three to nine digits must be repeated by the subject. Then additional sets of from two to eight digits must be repeated in reverse order.

Vocabulary. Forty words of increasing difficulty are presented both orally and visually. The words are defined orally by the subject.

Performance Subtests

Digit Symbol. This is a number-symbol association task in which unique symbols are paired with the numerals, 1-9. Subjects are required to write appropriate symbols in boxes below 90 presentations of numerals.

Picture Completion. A missing part in each of 21 pictures must be identified by the subject. This is the only Performance subtest requiring no manipulation of test materials by the subject, although he may respond by either verbally naming the missing element or by pointing to where the element should be on the picture.

Block Design. In the WAIS adaptation of the Kohs Block Design Test, the subject must reproduce ten designs using blocks that have two red, two white, and two half-red, half-white sides. Four blocks are used for the first six designs, and nine blocks are needed for the remaining items.

Picture Arrangement. Eight scrambled sequences of pictures are presented to the subject who must put each sequence in the proper order. A sequence resembles a newspaper comic strip.

Object Assembly. On the last subtest, four cut-up figures are reassembled by the subject. Essentially, this task is similar to solving simple jigsaw puzzles without prior knowledge of what the completed figure will look like.

APPENDIX C

ADDITIONAL TABLES REFERENCED IN THE TEXT

TABLE 1-1

Correlation Coefficients Among Session A Cognitive Measures for the Total Sample

	VOC	VMI	TH	MR	SAM	LAD	WBE	CR	VMD	PED	HP	PFB	NC	SPV	PMS
VOC	1.0000	.1174	.4257	.2362	.2716	.2243	.4787	.1853	.1635	.5456	.3474	.2255	.2473	.4555	.4325
VMI	.1174	1.0000	.0728	.0679	.0922	.1496	.1533	.1420	.4591	.2080	.2095	.0557	.1716	.1464	.1751
TH	.4257	.0728	1.0000	.2858	.1609	.2093	.4057	.2368	.0293	.2599	.2451	.3064	.1213	.3022	.3662
MR	.2362	.0679	.2858	1.0000	.1581	.3417	.1891	.6129	.1094	.2638	.4187	.4079	.0935	.2491	.5002
SAM	.2716	.0922	.1609	.1581	1.0000	.2676	.3177	.2872	.1289	.4326	.4098	.2208	.5081	.1650	.3239
LAD	.2243	.1496	.2093	.3417	.2676	1.0000	.2262	.3549	.1431	.3125	.3601	.3132	.2300	.1804	.3538
WBE	.4787	.1533	.4057	.1891	.3177	.2262	1.0000	.2790	.1733	.3892	.3200	.2507	.2672	.3149	.3627
CR	.1853	.1420	.2368	.6124	.2872	.3549	.2790	1.0000	.1005	.3146	.4716	.4881	.1904	.2626	.4142
VMD	.1635	.4591	.0293	.1094	.1289	.1431	.1733	.1005	1.0000	.2290	.1996	.0415	.2191	.1150	.1441
PED	.5456	.2080	.2599	.2638	.4326	.3125	.3892	.3146	.2290	1.0000	.4696	.2999	.4772	.4444	.4797
HP	.3474	.2095	.2451	.4187	.4098	.3601	.3200	.4716	.1996	.4696	1.0000	.4258	.3098	.3162	.4970
PFB	.2255	.0557	.3064	.4079	.2208	.3132	.2507	.4881	.0415	.2999	.4258	1.0000	.1444	.2352	.4473
NC	.2473	.1716	.1213	.0935	.5081	.2300	.2672	.1904	.2190	.4772	.3098	.1445	1.0000	.2430	.2788
SPV	.4555	.1464	.3022	.2490	.1650	.1804	.3149	.2628	.1150	.4444	.3162	.2352	.2430	1.0000	.4336
PMS	.4325	.1751	.3662	.5002	.3239	.3538	.3627	.4142	.1441	.4797	.4970	.4473	.2788	.4336	1.0000

TABLE 1-2

Unrotated Principal Component[a] Loadings
of Session A Tests for the Total Sample

Variable	Factor 1[b]	Factor 2	Factor 3	Factor 4
VOC	.6369	.1763	-.4646	.1980
VMI	.3035	.4697	.4697	.4369
TH	.5172	-.1632	-.4088	.3027
MR	.5910	-.5002	.2578	.1181
SAM	.5532	.2484	.0132	-.5839
LAD	.5382	-.1258	.2521	-.1102
WBE	.5976	.1570	-.3211	.1223
CR	.6353	-.4091	.3258	-.0462
VMD	.3072	.5289	.4585	.3574
PED	.7274	.2735	-.1390	-.1310
HP	.7161	-.0653	.1908	-.1226
PFB	.5835	-.4263	.1156	-.0284
NC	.5083	.4518	.0225	-.4811
SPV	.5742	.0705	-.3102	.2338
PMS	.7481	-.1529	-.0304	.0719
Eigenvalue	5.0954	1.5710	1.3199	1.1578
Percent of Total Variance	34.0	10.5	8.8	7.7
Percent of Common Variance	55.8	17.2	14.4	12.6

[a] Communalities of one were used to generate the factor matrix.

[b] Factor score coefficients from the first principal component were used to generate factor scores.

TABLE 1-3

Varimax Factor Loadings of Session A Tests for the Total Sample

Variable	Spatial Factor	Verbal Factor	Speed Factor	Memory Factor	Estimated Communality
VOC	.0834	.7978	.2078	.0718	.6918
VMI	.0878	.0867	.0360	.8412	.7242
TH	.2452	.6932	-.0868	-.0689	.5529
MR	.8031	.1700	-.0660	.0408	.6799
SAM	.2142	.1141	.8058	-.0262	.7089
LAD	.5349	.0912	.2639	.1309	.3812
WBE	.1356	.6435	.2409	.0967	.4998
CR	.8074	.0938	.1229	.0595	.6793
VMD	.0496	.0710	.1261	.8298	.7120
PED	.2302	.5073	.5489	.1697	.6404
HP	.5828	.2412	.3797	.1628	.5685
PFB	.6929	.2120	.0750	-.0750	.5363
NC	.0576	.1443	.8061	.1430	.6944
SPV	.1775	.6574	.1046	.1040	.4855
PMS	.5569	.4805	.2007	.0887	.5892
Percent of Total Variance	19.6	17.6	13.5	10.2	61.0
Percent of Common Variance	32.2	28.9	22.2	16.7	100.0

TABLE 1-4

Correlation Coefficients Among Session A Cognitive Measures for the AEA Sample

	VOC	VMI	TH	MR	SAM	LAD	WBE	CR	VMD	PED	HP	PFB	NC	SPV	PMS
VOC	1.0000	.1491	.4961	.2312	.3091	.2605	.5029	.1661	.1352	.6080	.3063	.2913	.3768	.5140	.5169
VMI	.1491	1.0000	.0983	.0612	.1735	.1299	.1193	.1426	.3802	.2025	.2235	.0182	.2059	.1462	.1713
TH	.4961	.0983	1.0000	.3363	.3329	.2222	.5175	.3014	-.0408	.4203	.3618	.2808	.2653	.3306	.4645
MR	.2312	.0612	.3363	1.0000	.2034	.3921	.3045	.6572	.0679	.2788	.4807	.4172	.1748	.2197	.5310
SAM	.3091	.1735	.3329	.2034	1.0000	.2401	.3604	.2363	.0900	.4555	.4057	.2156	.5565	.2302	.4101
LAD	.2605	.1299	.2222	.3921	.2401	1.0000	.3070	.3786	.0883	.3450	.4218	.2678	.2608	.1209	.4078
WBE	.5029	.1193	.5175	.3045	.3604	.3070	1.0000	.3537	.1160	.4425	.4039	.2826	.4166	.3505	.4410
CR	.1661	.1426	.3014	.6572	.2363	.3786	.3537	1.0000	.0563	.3232	.4846	.4889	.2396	.1813	.4226
VMD	.1352	.3802	-.0408	.0679	.0900	.0883	.1160	.0563	1.0000	.1822	.1638	.0420	.2733	.1141	.1359
PED	.6080	.2025	.4203	.2788	.4555	.3450	.4425	.3232	.1822	1.0000	.4626	.3358	.5173	.4191	.5308
HP	.3063	.2236	.3618	.4807	.4057	.4218	.4039	.4846	.1638	.4626	1.0000	.4319	.3674	.2406	.5631
PFB	.2913	.0182	.2808	.4172	.2156	.2678	.2826	.4889	.0420	.3358	.4319	1.0000	.2057	.1886	.4941
NC	.3768	.2059	.2653	.1748	.5565	.2608	.4165	.2396	.2733	.5173	.3674	.2057	1.0000	.3100	.3895
SPV	.5140	.1462	.3306	.2197	.2302	.1209	.3505	.1813	.1141	.4191	.2406	.1886	.3100	1.0000	.4423
PMS	.5169	.1713	.4645	.5310	.4101	.4078	.4410	.4225	.1359	.5308	.5631	.4941	.3895	.4423	1.0000

TABLE 1-5

Unrotated Principal Component[a] Loadings of Session A Tests for the AEA Sample

Variable	Factor 1[b]	Factor 2	Factor 3	Factor 4
VOC	.6726	.3189	-.3480	.2508
VMI	.2846	.3751	.5866	.2460
TH	.6291	.0054	-.3895	.0830
MR	.6130	-.5399	.1435	.1525
SAM	.5900	.2410	.0026	-.5827
LAD	.5415	-.2418	.2072	-.1575
WBE	.6793	.1075	-.2098	-.0171
CR	.6148	-.5187	.2274	.0370
VMD	.2300	.4176	.6409	.2555
PED	.7462	.2525	-.0945	-.0578
HP	.7190	-.1889	.2200	-.1208
PFB	.5704	-.4048	.0077	.0813
NC	.6150	.3917	.1055	-.4121
SPV	.5380	.2951	-.2860	.4165
PMS	.7985	-.0957	-.0467	.1099
Eigenvalue	5.5430	1.6184	1.3395	.9649
Percent of Total Variance	37.0	10.8	8.9	6.4
Percent of Common Variance	58.7	17.1	14.0	10.2

a. Communalities of one were used to generate the factor matrix.

b. Factor score coefficients from the first component were used to generate factor scores.

TABLE 1-6

Varimax Factor Loadings of Session A Tests for the AEA Sample

Variable	Spatial Factor	Verbal Factor	Speed Factor	Memory Factor	Estimated Communality
VOC	.1076	.8238	.1998	.0892	.7381
VMI	.0803	.0857	.1004	.7762	.6263
TH	.2876	.6366	.2012	-.1609	.5544
MR	.8245	.1738	-.0261	.0195	.7111
SAM	.1631	.1731	.8302	.0067	.7458
LAD	.5634	.0569	.3026	.0846	.4194
WBE	.2888	.5564	.3525	.0001	.5173
CR	.8270	.0830	.0809	.0532	.7001
VMD	.0307	.0415	.0860	.8327	.7034
PED	.2548	.5640	.4750	.1557	.6329
HP	.6417	.1918	.3742	.1643	.6156
PFB	.6584	.2379	.0524	-.0561	.4960
NC	.1069	.2602	.7634	.2254	.7126
SPV	.0667	.7750	-.0012	.1635	.6319
PMS	.5631	.5253	.2415	.0986	.6610
Percent of Total Variance	20.8	18.9	13.3	10.0	63.1
Percent of Common Variance	33.0	30.0	21.1	15.9	100.0

TABLE 2-1

Intercorrelations[a,b] of Session A Tests with WAIS Subtests and Scales

Session A Tests	WAIS Subtests											WAIS Scales		
	I	C	A	S	D	V	DS	PC	BD	PA	OA	VIQ	PIQ	FIQ
VOC	.5716	.4410	.3835	.4254	.3240	.7348	.2685	.2688	.2628	.3041	.1703	.6313	.3679	.6044
VMI	.1732	.1223	.1235	.0737	.1025	.1284	.1558	.1722	.1122	.1518	.1445	.1652	.2186	.2159
TH	.4488	.3374	.3010	.3554	.2626	.4486	.1347	.2746	.2746	.2805	.2574	.4939	.3676	.5094
MR	.3086	.2401	.3820	.2070	.2101	.2552	.0808	.3844	.4985	.2513	.3807	.3690	.4904	.4819
SAM	.2296	.0327	.4461	.1263	.2361	.1497	.5664	.0941	.2509	.1942	.0480	.2748	.3433	.3491
LAD	.2198	.0950	.3006	.1356	.2502	.1562	.2670	.2838	.3110	.2254	.2000	.2696	.3838	.3681
WBE	.3657	.1880	.3125	.2718	.3465	.3685	.2444	.1876	.2270	.1709	.1529	.4226	.2963	.4305
CR	.2377	.1602	.3651	.2007	.2284	.1697	.2352	.3362	.4985	.1732	.3632	.3168	.5042	.4534
VMD	.1662	.1525	.1006	.0867	.0649	.1226	.1813	.1434	.1325	.1008	.1223	.1593	.2165	.2106
PED	.4476	.2932	.4311	.3335	.2940	.4303	.4712	.2813	.3176	.3505	.2653	.4995	.4982	.5731
HP	.2993	.1303	.3321	.2274	.2150	.2551	.4119	.3297	.3905	.2587	.2669	.3208	.5062	.4586
PFB	.2367	.1668	.3046	.2259	.1526	.2314	.2270	.2815	.4923	.2621	.4917	.3012	.5216	.4538
NC	.1260	.0770	.2306	.1316	.1109	.1060	.6042	.0786	.1436	.1942	.0731	.1818	.3248	.2777
SPV	.4068	.3670	.2786	.3546	.1722	.4594	.2279	.2768	.2578	.3291	.2328	.4592	.3768	.4908
PMS	.4359	.3203	.4837	.3291	.2584	.4090	.2779	.3883	.5178	.3662	.4047	.5075	.5756	.6230

a. A correlation greater than .08 is significant at the 5% level; one greater than .10 at the 1% level; and one greater than .14 at the .1% level.

b. Each correlation was computed from data on 448 individuals.