UNIVERSITY OF ALBERTA

Theoretical and Factorial Conceptualizations of Human Intelligence: An Evaluation of the Stanford-Binet Intelligence Test Fourth Edition

BY

(C) Richard Spelliscy

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF Doctor of Philosophy

IN

Counselling Psychology

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

EDMONTON, ALBERTA

SPRING, 1991

ISBN 0-315-66764-8

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Richard Spelliscy

TITLE OF THESIS Theoretical and Factorial Conceptualizations of Human Intelligence: An Evaluation of the Stanford-Binet Intelligence Test Fourth Edition

DEGREE: Doctor of Philosophy

YEAR THIS DEGREE GRANTED: 1991

R J Spelliscy

Richard Spelliscy

9818 - 92 Avenue

Edmonton, AB T6E 2V5

Date: March 21, 1991

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommends to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **THEORETICAL AND FACTORIAL CONCEPTUALIZATIONS OF HUMAN INTELLIGENCE: AN EVALUATION OF THE STANFORD-BINET INTELLIGENCE TEST FOURTH EDITION** submitted by **RICHARD SPELLISCY** in partial fulfillment of the requirements for the degree of **DOCTOR OF PHILOSOPHY** in **COUNSELLING PSYCHOLOGY.**

Dr. H.L. Janzen, Supervisor

Dr. J.G. Paterson

Dr. F.J. Boersma

Dr. E.L. Eberlein

Dr. K.L. Ward

Dr. Reddon

Dr. Anita Li

Date: Mar 19/91

DEDICATION

To
Denise Perret
for her friendship and
faith in me

ABSTRACT

Theoretical and factorial conceptions of human intelligence have been mainstays of psychological research since the turn of the century. The early work of Alfred Binet and the development of factor analysis are identified as key elements in shaping subsequent theoretical and psychometric trends. The Fourth Edition of the Stanford-Binet is influenced by developments in both of these areas through the adoption of a hierarchical model of intelligence.

The theoretical and empirical basis of this instrument is examined by a review of the development and evolution of the Binet scales, major theories of intelligence and existing factorial studies. The Stanford-Binet protocols of 371 Education clinic clients between the ages of 3 and 23 years were analyzed using both exploratory and confirmatory factor analytic techniques. These results were compared to the test authors' technical data.

The three level hierarchical model espoused by Thorndike, Hagen and Sattler (1986) was not supported by the Clinic data. Modest support was found for both developmental and modified test models. On the basis of these results it is tentatively concluded that a two tier model with a General factor at the highest level and Verbal, Abstract/visual, Quantitative and Short term memory factors at the lower level, under certain restrictions, best describes what the Stanford-Binet Fourth Edition measures.

Test users, as a result, are cautioned against the strict implementation of Stanford-Binet Fourth Edition theoretical model. Ethical and practical considerations in the interpretation of this instrument were provided. Suggestions for future research were made.

ACKNOWLEDGEMENTS

I wish to thank the many clients, support personnel, graduate students, and faculty associated with the University of Alberta, Education Clinic for participating in this study. I am especially indebted to Laura Clarke, Bruce Fritzke, and Rivy Meloff who freely shared both research and ideas that contributed greatly to the scope of this study. I gratefully acknowledge the invaluable contributions of my supervisory committee members Drs. Frederic Boersma, Henry Janzen, and John Paterson. Dr. Janzen, my supervisory committee Chairman, deserves special mention for his ongoing guidance, instruction and support, for which I sincerely appreciate.

Drs. Kenneth Ward, Larry Eberlein, and John Reddon as members of my committee each had an immeasurable influence. Dr. Ward's enthusiastic willingness to participate on my committee was greatly appreciated. Dr. Eberlein's expertise and commitment to ethical issues provided an ongoing reminder of the importance of the task at hand. Dr. Reddon deserves special acknowledgement for his excellent research skills and the ongoing consultation provided by him. Without his extensive knowledge and willingness to freely give his time this study would not have been possible. I would also like to thank Ms. Vicki Ross for her friendliness, perseverance, and willingness to type the numerous tables and drafts.

TABLE OF CONTENTS

CHAPTER PAGE

CHAPTER	PAGE

CHAPTER PAGE

CHAPTER PAGE

LIST OF TABLES

TABLE PAGE

TABLE		PAGE

TABLE PAGE

TABLE PAGE

TABLE PAGE

TABLE PAGE

LIST OF FIGURES

CHAPTER 1

INTRODUCTION TO THE STUDY OF HUMAN ABILITIES

The question that factor analysis of human ability attempt to answer "has baffled psychologists and philosophers of mind for centuries" (Royce, 1958, p. 141). Like many great problems one can trace the earliest writings on the nature of human intelligence to the Greek philosophers. Plato (429-348 B.C.) compared human intellect to a charioteer citing the powers of perception and control as being salient. He also introduced the term 'nous' or intellect (Cattell, 1987). Plato's student Aristotle (384-322 B.C.) is said to have made the distinction between 'orexis' emotional and moral functions and 'dionia' cognitive and intellectual functions (Burt, 1955).

These early theories were subsequently criticized on both theoretical and physiological grounds. Theoretically, they were viewed as being overly simplistic and prone to discussing the mind as if it were an object. Physiologically, they were viewed as being naive by "considering the brain a sort of sponge radiator for cooling the blood rather than the seat of intelligence, which they placed elsewhere" (Cattell, 1987, p. 2).

The early Greek theories were profoundly influenced by emerging Christian thought. The "very father of faith", St. Augustine (354-430), considered by many to be the first psychologist, proposed "a three-fold classification of mind as intellect, will and self-conscious memory..." (Royce, 1958, p. 141). During the middle ages with the rise of scholasticism that St. Thomas Aquinas (1225-1274) described such intellect as it is more commonly known today as intelligence or understanding. At approximately the same time King Edward 1 (1239-1307) enacted a progressive law distinguishing between those born lacking intelligence and those who possess intelligence but because of emotional or physical impairment have lost or diminished abilities (Cattell, 1987).

It was not, however, until the nineteenth century that the concern for the mentally and emotionally handicapped re-emerged. The accurate differentiation between the intellectually disadvantaged and emotionally impaired became, central. A concommitant concern arose in the education system with reference to the differentiation of learning abilities (Anastasi, 1982, 1986). "But even by the mid-nineteenth century there had been little advanced in fundamental concepts about the nature of that intelligence which had failed so mysteriously in the mental defective" (Cattell, 1987, p. 2). Parallel to this slowness of mankind in viewing themselves as appropriate objects of scientific study was the lack of appropriate statistical methods. DuBois (1970) provides an extensive review of the history of psychological testing. Nunnally (1978) stated:

> the failure to perform systematic studies of individual differences may have been do to the lack of adequate methods of statistical analysis - correlational analysis, factor analysis, and related methods of multivariate analysis. Although many forms of mathematics had reached a high level of development by 1800, it was not until over 100 years later that statistical methods were developed that would be required to construct psychometrically sound tests of ability and to investigate their inner relationships (p. 502).

Sir Francis Galton - British biologist, psychologist and statistician is generally regarded as the father of mental tests (Boring, 1923). Influenced by the earlier work of his cousin Charles Darwin on the differential survival abilities of plants and animals, Galton became interested in the range of human abilities particularly as it relates to heritability of these traits. He coined the term 'mental test' to describe his attempts to measure human abilities. However, under the influence of early German psychology he chose to focus on a variety of physical and sensory measurements, as he believed that "the only information that reaches us concerning outward events appears to pass through the avenue of our senses; and the more perceptive the senses are of difference, the larger is the field upon which our judgement and intelligence can get" (Galton, 1883, p. 27)(also see Nunnally, 1978; Tuddenham, 1962). Nunnally (1978) has summarized Galton's main contribution to the contemporary study of human abilities:

> To analyze the obtained data Galton made use of statistical methods, and with these he determined averages and measures of dispersion. He particularly needed a measure of association, or correlation, to determine the amount of resemblance between the characteristics of fathers and their sons. For this purpose, he made the first steps in the development of correlational analysis (p. 503).

One of Galton's students Karl Pearson extended his predecessor's work in the mathematical measurement of human abilities and became an eminent statistician in his own right. Pearson was responsible for the chi- square distribution. He also derived the product moment correlation coefficient, multiple and partial correlations; the basic foundations for factor analytic techniques (i.e., principal components) in 1901. Pearson also founded Biometrika which is a highly influential statistical journal today. Pearson, his son, Ergon, and collaborator J. Neyman and R. Fisher have had a substantial impact on statistics as they are known today.

Pearson, however, despite these great contributions of mathematical genius and like many of the great researchers of his time (including Binet, Titchner, Ebbinghaus, Woodsworth, and Cattell) was unable to formulate a method of identification of common factors of human ability despite the belief on their possible existence. Royce (1958) writes:

> It was not until 1904, in an epoch-making paper, that Charles Spearman promulgated his now famous two factor theory of intelligence: "All branches of intellectual activity have in common one fundamental function (or group of functions), whereas the remaining or specific elements of the activity seem in every case to be wholly different from that in all the other cases" (1904, p. 284).

Approximately at the same time as Spearman's early works and just prior to the discovery of his "two-factor theory" Alfred Binet was pursing a variety of areas in physiological psychology. The number of topics he explored was, indeed, phenomenal ranging from the psychic life of microorganisms (1887) and insects (1894) to that of great mathematicians and skilled games players (1894) to that of anthropometrics (1910) to projective testing (1896 and 1906)(Tuddenham, 1962). Peterson (1925) remarked "through all this search, Binet has shown a master's hand in discovering realities in human nature and in "*letting facts lead*" (p. 149). Consequently, Binet's research was relatively more pragmatic than his psychological contemporaries in Britain and Germany (Tuddenham, 1962).

<u>The Development of the Stanford-Binet Scales</u>

According to Dunn (1980) Binet's trial and error approach was timely and advantageous. He identified two events as being important to Binet's historic course. The first was Binet's increasing awareness that he should disregard the measurement of physical or elementary units of mental activity, identified by narrow and separate measures, and focus on the concept of general intelligence as initially recognized by Galton and that would manifest itself in a global mental ability score. This view provided him with a theoretical foundation to explore a wide variety of interrelated tests. The second direct event according to Dunn was the learning of a set of test items, developed by two fellow Parisians, that fit this model. Dunn (1980) writes:

> These had been developed by ... "Dr. Blin, and his student, Dr. Damaye, "the unsung heroes" in this case. For both his ideas and his items, Binet drew heavily from the work of these men and others and gave them due credit...In no way is this intended to disparge the extensive work done by Binet and Simon in refining the Blin items, sorting those that contributed the most to a global measure of intelligence, finding and devising others, and arranging the array in order of difficulty"
> (p. v).

A final concommitant impetus to what is generally considered to be the first major test of human intelligence was the French Ministry for Public Instruction's directive that, "no child suspected of retardation should be eliminated from the ordinary school and admitted into a special class, without first being subjected to a pedogogical and medical examination from which it could be certified that because of the state of his intelligence, was unable to profit, in an average measure, from the instruction given him in the ordinary schools" (Binet & Simon, 1905a, pp. 163-164).

By this time Binet and his colleagues had been studying individual differences for approximately fifteen years and they believed that all present intellectual classifications were lacking. Once again, they turned to the work of Blin and Damaye whose study of 250 idiots, imbeciles, and morons, they described as the "first attempt to apply a scientific method to the diagnosis of mental ability" (Binet & Simon, 1905a, p. 28). Recognizing the contributions of Blin and Demaye and their own experiences Binet and Simon came to the conclusion that what was required to make meaningful conclusions about individual differences in abilities was a variety of easily measured tasks that would display age-grade relationships. The importance of norms as benchmarks for "normal" development had been established in their earlier anthropometric studies. Binet and Simon were also aware of the need for standardization in both presentation and scoring of responses and as well as the need for efficient measures (Wolf, 1973).

Approximately one year later, the 1905 scale was introduced. This scale was viewed as the "first successful test of general intelligence, the direct ancestor of the Stanford-Binet and most other modern intelligence scales... (and as noted) preconceptions and theories about the nature of general intelligence entered in hardly at all" (Tuddenham, 1962, p. 483). Binet (1905b) wrote:

> The fundamental idea of this method is the establishment of what we shall call a "metrical scale of intelligence"; this scale is composed of a series of tests, of increasing difficulty starting at one end from the lowest intellectual level that can be observed, and emerging at the other at the level of average, normal intelligence, with each test corresponding to a different mental level (p.194).

Consequently, test items for the 1905 scale and all subsequent scales excluding the Stanford-Binet (Fourth Edition) were selected for their ability to differentiate levels of mental ability (i.e., metric properties) rather than a preconceived model or definition of intelligence. Binet and Simon's (1905b) view of intelligence had, however, solidified. They saw it as consisting of a single fundamental faculty. "This faculty is judgment, otherwise called good sense, initiative, the faculty of adapting one's self to circumstances. To judge well, to comprehend well, to reason well, these are the essential activities of intelligence" (p. 42-43). Three other words complete Binet and Simon's definition of intelligence. These are invention (i.e., solution) which is made after one comprehends; direction, that is maintaining sustained action after one knows what to do and; censure or self criticism. For Binet and Simon intelligence consists of these four words; comprehension, invention, direction and censure (Wolf, 1969b). Such a view emphasizes both the active and structural components of intelligence. It has been argued, however, that their initial measure and subsequent revisions up until the Stanford-Binet (Fourth Edition) do not "adequately reflect the breadth of the construct of intelligence even as defined by themselves" (Yeboah & Peat, 1986, p. 9). Specifically, intelligence is defined in very broad terms, measured by a relatively narrow range of tasks and relagated a unitary status (Yeboah & Peat, 1986).

The Stanford-Binet: (Fourth Edition)

The Stanford-Binet (Fourth Edition) (hereafter referred to as the SBIV) represents a basic theoretical and empirical restructuring of all earlier revisions and attempts to address criticisms put forth by such researchers as Yeboah and Peat (1986) and others." Beginning with a hierarchical model of cognitive abilities, the test construction process (spanning some 8 years) pursued the dual goal of retaining as many item types as possible from the earlier editions while incorporating current ability constructs" (Anastasi, 1988, p. 200) . As noted earlier the authors of the SBIV (Thorndike, Hagen, & Sattler, 1986) determined that four areas of cognitive ability should be evaluated: verbal reasoning, quantitative reasoning, abstract/visual reasoning, and short-term memory. They also decided that the SBIV should continue to provide a global index of functioning that would represent what is commonly known as *'g'* or general reasoning ability. These separate areas were then united in a hierarchical view of intelligence which provided their theoretical model of human intelligence (see Figure 1.0).

To establish the construct validity of their model, Thorndike, Hagen and Sattler (1986b) utilized an unnamed variant of confirmatory factor analysis. The authors concluded that "confirmation was obtained for each of the four Area scores, though the group factor loadings are in all cases lower than the general factor loadings" (p. 54). In an earlier statement in the same manual they wrote "still, the general ability factor, *'g '*, refuses to die. Like a phoenix, it keeps rising from its ashes and will no doubt continue to be an enduring part of psychometric theory and psychometric practice" (p. 6).

Several writers have questioned Thorndike, Hagen and Sattler's (1986b) interpretation of their factorial results (Anastasi, 1988b; Osberg, 1986; Reynolds, 1988; Sandoval & Irvin, 1988; and Vernon, 1987). Sandoval and Irvin (1988) summarize an alternative view of the test authors' findings.

FIGURE 1.0

Theoretical Model of the Stanford-Binet: Fourth Edition

Has been removed because of the unavailability of copyright permission.

Thorndike, Hagen and Sattler (1986b, p. 9)

> The results revealed relatively low loadings between the tests in the battery and the factor they were expected to be associated with. The values reported, for example, are lower than those reported by Kaufman & Kaufman (1983) on the Kaufman Assessment Battery for Children (K-ABC) in a similar analysis. Instead, it appears that each of the tests has a relatively high degree of specificity and error, and that each measures something different from the other subtests in its assigned area. Thus, there may be support for the total composite score as a measure of 'g', but as yet there is no evidence supporting the use of the area scores. The results of our and other's (Reynold's personal communication, December 15, 1986) exploratory factor analysis suggest that the tests load on fewer than four factors at any age and that different factors emerge at different ages. The test probably does not measure the same thing at different ages (p. 160).

These critics do agree that the SBIV appears to possess satisfactory evidence for the construct validity of its composite index. They are unanimous in their view that the usefulness of the four cognitive area scores has yet to be determined. Indeed, several recommend the use of alternative tests (Cronbach, 1988; Sandoval & Irvin, 1988) and one even suggests the continued use of the earlier Stanford-Binet revision (Vernon, 1987). Thus, the underlying theoretical and empirical model of the SBIV is seriously being questioned and as a result, the need for its further evaluation is critical.

It is also important to evaluate the factor structure of the SBIV for practical reasons. Thorndike, Hagen and Sattler (1986) correctly note that the results of this "test are most frequently used, together with other information, to make recommendations for educational intervention..." (p. 9). The four cognitive ability areas are seemingly proposed to provide the clinician with valuable diagnostic and remedial information in these areas. The utility of any score rests on its verification. Clearly, if its existence is uncertain it is imperative that its use be discontinued or at the very least appropriate qualifications should be made. The widespread use of the Binet scales, rivaled only by the Wechsler scales serves only to increase the urgency of such evaluation.

The existing empirical literature on the SBIV is small and the majority of articles are of a critical review nature. Although the '*Technical Manual*' for SBIV provides substantial empirical information its factorial structure, as noted above, is questionable. Independent research on populations different from those involved in original test development is standard psychometric practice and ideally takes place prior to the wide spread release of new instruments.

The early Binet scales (Binet & Simon, 1905, 1908, 1911; Terman & Merrill, 1937, 1960 and 1973) provide important historical and scientific information on the evolution of the SBIV. They highlight the important contribution to the field of psychometrics made by Binet and subsequently Terman and his associates at Stanford University. They also allow one to identify both the consistencies and innovations in the SBIV. Thus, one can assess as both Vernon (1987) and Sandoval and Irvin (1988) have suggested whether or not "the test resembles a new test more than a revision of the old instrument" (p. 157).

A review of the major alternative factorial theories of intelligence that coexisted with the development of the Binet scales is also viewed as essential. It is necessary to the evaluation of the theoretical model underlying the SBIV. The proposed theoretical structure borrows extensively from at least two theorists (i.e. Cattell, 1971; and Vernon, 1961).

Finally, existing factorial studies of the earlier Binet scales are considered. This is meaningful as these studies have suggested the presence of group factors. It is important to determine if these group factors are comparable to those proposed by the SBIV. In summation, the purpose of this study is to examine the internal structure of the SBIV to determine if it conforms to the model hypothesized by Thorndike, Hagen and Sattler (1986b). Hence, it is an attempt to make a definitive contribution to one type of validity. This is construct validity.

The Binet scales and subsequent revisions have and continue to be a hallmark of psychometric practice. It is important to determine if the SBIV is worthy of this honor. All findings and conclusions are, however, subject to the parameters surrounding this study.

Plan of Dissertation

Five main topics are addressed in (Chapter 2) the abbreviated literature review. First, the early Binet-Simon scales and subsequent American revisions leading to the SBIV are examined. Next, the major factor analytic theories of intelligence are presented along with their relationship to the SBIV. Third, existing factor analytic studies of the Binet scales are examined to determine if consistent factorial groupings are evident. The construction of the SBIV is then reviewed. Finally, pertinent and salient issues in factor analytic research are highlighted.

Chapter 3 outlines the research design and procedures that are used. Descriptions of the instrument, participants, procedures, and statistical techniques employed for handling missing data are provided. The limitations of this research are then presented.

The research objectives and questions are then stated. Chapter 4 is a presentation of the results. The relationship of these results to the research objectives and practical considerations in using the SBIV are examined in Chapter 5. Recommendations for future research are also provided in Chapter 5.

Study Limitations

McNemar (1942), Guilford (1954) and Gorsuch (1983) have addressed the most common faults of factor analytic studies. The most relevant to the proposed study, at this time, is the size and nature of the sample. The proposed sample size is adequate according to recent literature (Baggaley, 1982; Loo, 1983). Reddon (Under Review, 1988) has determined that the sample size requirements reported by Baggaley (1982) are over-estimates. A larger sample however, would likely provide more consistent findings.

The second sample concern is its potentially idiosyncratic nature. Although, all three key ability levels are adequately represented, a considerable proportion of the sample was referred because of learning difficulties. Guilford (1952), Sternberg (1977), and Comrey (1978) have stated that homogeneous sampling tends to lower correlations. On the other hand, diverse samples that are pooled together merely to increase the sample size can lead to problems in the opposite direction. The proposed sample is perceived to be adequate given; the logistical effort required to obtain completed protocols, the urgent need to examine the factorial structure of the SBIV and the value of obtaining a greater understanding of this test when it is used with clinic populations. Appropriate cautions are, however, required and duly specified.

Research Objectives

1. To provide descriptive statistics including means, ranges, standard deviations and correlations for the SBIV subtest, area and composite scores.

2. To factor analyze the SBIV according to the three age levels identified by Thorndike, Hagen & Sattler (1986) to determine if the obtained structure compares favorably to that proposed by its authors.

3. To compare the different factor analytic methodologies to determine if similar results are obtained and if one solution, from a psychological perspective, provides more easily understood results.

4. To offer some suggestions to clinicians who may wish to utilize the four area scores of Verbal Reasoning, Abstract/visual Reasoning, Quantitative Reasoning and Short-Term memory.

5. To provide suggestions for future research on the SBIV.

Research Questions

1. Will the factorial results from the clinic sample conform to the theoretical, hierarchical model proposed by Thorndike, Hagen & Sattler (1986)? More specifically, is a large common factor found with two or more specific group factors?

2. Will the factorial results from the clinic sample conform to the factor structure reported in the SBIV *Technical Manual*? That is: At the 2 to 6 years of age level are two small group factors (i.e. Verbal and Abstract/visual) found in addition to a large common factor; At the 7 to 11 year level are three small group factors (i.e. Verbal, Memory, Abstract/visual) in addition to a large common factor present?; At the 12 to 23 year level are four small group factors (Verbal, Abstract/visual, Memory and Quantitative) in addition to a large common factor present?

3. Do the obtained factor analytic results resemble those that were found from factor studies of earlier Stanford-Binet editions?

4. If the factorial results do not support the hypothesized structure, what is the nature of the construct or constructs underlying the SBIV?

CHAPTER 2

LITERATURE REVIEW

Historical Background to the Stanford-Binet (Fourth Edition)

The Early Binet-Simon Scales

The 1905 Binet-Simon Scale was viewed as an important scientific breakthrough and it has had a sizable influence on subsequent developments in psychology and pedagogy (Wolf, 1973). The scale was guided by 15 years of empirical experimentation, sparked by the work of Blin and Damaye and requested by the French Minister of Public Instruction to assist in the differentiation of student abilities. Binet and Simon (1905b) stated their purpose was "to make all our tests simple, rapid, convenient, precise, heterogeneous, holding the subject in continued contact with the experimenter, and bearing principally upon the faculty of judgment" (p. 41). The final thirty tests making up the 1905 scale included items in the areas of visual perception, abstract/spatial reasoning, auditory/visual memory, social comprehension, vocabulary and language development.

The 1908 revision of the Binet-Simon scale was an attempt to correct many of the weaknesses that were perceived in the original scale. A larger and more representative sample of children was used to enhance the accuracy of age related items. The total number of test items was increased to 59 and the age range was broadened. Importantly, this work resulted in the development of the first age scale. Using the symmetrical bell shaped curve as a criterion Binet and Simon placed an item at the year level where it was passed by two-thirds to three-fourths of the age related sample. Freeman (1955) has summarized the steps involved in the calculation of Mental Ages (MA):

> The mental age within the 1908 scale was found this way: the subject was credited with the age level at which he passed all the tests. To this basic level (now called the 'basal year') an additional year's credit was added for every five tests passed at higher levels. The total was the subject's mental age ... (p. 109).

The 1911 test revision saw a continued rearrangement of items (movement of items up and down the age scale), elimination of items that were redundant or too scholastic in nature and the addition of new items that passed experimentation. The age level was increased to extend from three years to adult level. Perhaps the most important change was the standardization of the number of tests at each age level. This allowed Binet to equate each additional test passed beyond one's chronological age with two-tenths of a mental year. Despite the explicit nature of his mental age concept, Binet refrained from adopting Stern's newly developed Intelligence Quotient (IQ). Binet rejected it on the grounds that differences in intelligence were qualitative as well as quantitative and that such an index may ignore the former (Thompson, 1984; Wolf, 1973).

In summation, Binet and Simon are generally credited for developing the first practical test of human intelligence (Matarazzo, 1972). Their early scales introduced many concepts familiar to modern day psychometricians including; global ability, age scale, mental age, basal and testing ceiling (French, 1986). The format of most major intellectual assessment tools today bear a strong resemblance to the early Binet-Simon tests. It is impressive to see how few of these early concepts or ideas have changed since their introduction.

The American Revisions

The Binet-Simon scales were quickly accepted by many countries and particularly by the United States. Although there were many American adaptations those developed at Stanford University were by far the most accepted and they soon prevailed. The first published revision by the Stanford group headed by Lewis Terman was in 1916.

The 1916 Stanford-Binet (as it came to be known) was an attempt to combine the features of the Binet scales with current psychometric knowledge. The 1916 revision increased the number of test items from 54 to ninety. The entire scale was restandardized on a sample consisting of 1400 Americans. Stern's concept of IQ was adopted. There were, however, many criticisms of this scale and several related to the adequacy of the standardization sample.

The second revision of the Stanford-Binet was published in 1937. According to Terman and Merrill (1953), the three main objectives of this revision were to (a) express test scores into age levels of performance, (b) measure problem solving skills by a variety of tasks that did not rely specifically on scholastic experiences, and (c) to determine general levels intelligence. To achieve these objectives the authors (a) obtained a sample that was more representative of the general United States population, (b) extended the scale downward to age two years and added three levels of superior performance, (c) more carefully standardized those age groupings previously identified as being weak, (d) increased the amount of nonverbal content, and (e) included a wider variety of items (Freeman, 1955; Terman & Merrill, 1937). The revision consisted of two equivalent forms (L and M) to allow for retesting to assess possible rater bias.

The purpose of the 1960 revision was to provide "test users with a single scale that, while preserving the characteristic features of previous revisions, eliminates out of date content and improves general structure" (Terman & Merrill, 1973, p. v). The best test items from Forms L and M were combined into single form.

The 1960 restandardization, however, was incomplete. Instead of selecting a new sample, data on 4,498 subjects who had been administered the test between 1950 and 1954 was used. According to Terman and Merrill (1973) the most important changes were structural and included replacing the 1937 IQ tables with Pinneau's Deviation IQ tables. These new tables adjust for differences in variability of IQs at specific age levels and allow for more accurate comparison of scores across age levels.

The 1960 Form L-M was restandardized in 1972. The purpose of the restandardization was simply to update norms and test content changes were intentionally kept to a minimum. The subjects were siblings of individuals who were selected for the Cognitive Abilities Test standardization sample. Subjects included minority groups and they were stratified according to ability levels, economic status and geographical area. The entire sample consisted of approximately 2100 subjects and an attempt was made to locate

100 subjects to each Stanford-Binet year level. Despite these efforts, subsequent reviews have criticized the sampling procedures employed (Davis & Rowland, 1974; Holroyd & Bickley, 1976; Salvia, Ysseldyke & Lee, 1975; Waddell, 1980).

The next restandardization was not to take place until some sixteen years later. Like its predecessors, the SBIV has generated both praise and criticism. Prior to reviewing this revision it is important to briefly state the theoretical and empirical bases of the SBIV.

Theoretical Background to the Stanford-Binet (Fourth Edition)

Spearman's Two Factor Theory of Intelligence

Whereas Binet assumed that a unitary or pervasive factor (i.e., judgment or adaptation) was the common denominator of human intelligence Spearman viewed it as a hypothesis yet to be tested (Brody & Brody, 1976; Nunnally, 1978). He proceeded to develop the mathematical criteria (i.e., factor analysis) to test for the existence of a general factor in a matrix of correlations among tests. A series of studies in 1904(a,b) and his main body of work "The Abilities of Man" published in 1927 outlined his findings. He found that the arrangement of human abilities could be expressed by a definite mathematical equation (i.e., tetrad) and summarized by his now famous two-factor theory of intelligence (also see Brody & Brody 1976; Butcher 1968; Nunnally 1978; Royce, 1958 and Tuddenham, 1962). Spearman stated:

> Whenever, the tetrad equation holds throughout any table of correlations and 'only' when it does so, then every individual measurement of every ability (or of any other variable that enters into the table) can be divided into two independent parts which possess the following momentous properties. The one part has been called the "general factor and denoted by the letter *g*; it has been named so because, although varying freely from one individual to individual, it remains the same for any one individual in respect of all the correlated abilities. The second part has been called the 'specific factor' and is denoted by the letter *s*. It not only varies from individual to individual, but even for any individual from each ability to another. The proof of this all important mathematical theorem has gradually evolved successive stages of completeness, and may now be regarded as complete. (1927, p. 74-75).

Although the empirical definition of g came rather quickly to Spearman the psychological interpretation proved much more difficult (Brody & Brody, 1976; Butcher, 1968). He reported that only one hypothesis appeared to fit the facts and concluded that g measures something analogous to energy, a concept originally put forth by Aristotle. Spearman's mental energy was thought to fuel the "engines" of the nervous system. Differences in ability were also determined by three qualitative laws (i.e., apprehension of experience, education of relations and eduction of correlates) and five quantitative ones (i.e., span, retentivity, fatigue, conation, and primordial potencies)(Brody & Brody 1976, Nunnally; 1978 and Tuddenham 1962).

Spearman's (1927) work gave tacit recognition to the existence of group factors. He maintained that specific factors were independent of each other except when task demands were very similar. Although he went on to say that cases of group factors were rare he saw four or five as warranting the most attention. These were the logical, psychological, mechanical, arithmetical and more tentatively, the musical (Brody & Brody, 1976; Butcher, 1968).

Spearman's work has continued to generate much criticism and praise. First, and not withstanding his own criticisms of Binet on this topic, Spearman's samples were small and nonrepresentative. Vernon (1961) states that this is one reason why group factors do not appear. Second, Spearman's decision to make each task substantially different often precludes the finding of group or second order factors. Fancher (1985b) and Vernon (1961) cite the work of British statistician Godfrey Thompson and stress that Spearman's two-factor theory was not the only possible explanation for the observed patterns of correlations. Fancher (1985a) recently checked and recalculated Spearman's original data. He discovered a large number of errors and concluded that while the essential patterns hold the relationships were considerably less perfect than Spearman claimed.

In summation, Spearman's main conclusions relating to the presence of g have proved to be sound and its presence in the theoretical model hypothesized for the SBIV can tentatively be accepted. Spearman's statements about the nature of intelligence were deduced from the types of measurements he proposed. He developed mathematical procedures that could be used to support or reject theories of human intelligence. Spearman also recognized that problem solving speed and intelligence were correlated. He did, however, disagree with Binet and Simon's theoretical position that their tests worked because they measured individually patterned intelligences. For Spearman, a general factor along with specific factors of different magnitudes explained intelligent behavior (Fancher, 1985,ab). Finally and perhaps most important of all, Spearman established an empirical tradition in the investigation of human abilities (Nunnally, 1978).

Thurstone's Primary Mental Abilities

Like Binet's ideas Spearman's concepts were quickly being studied in the United States. L.L. Thurstone (1938, 1941) turned around Spearman's question on whether or not the correlations in a matrix can be explained by a single overriding factor to "*how many* and *what kinds* of factors are needed to account for the observed correlations among tests of ability" (Nunnally, 1978, p 509).

Thurstone's work started with several theoretical assumptions. First, he believed that performance on any ability test was the result of a specific number of basic or primary abilities (a term used to describe psychological characteristics that are represented by a group factor). Second, that the number of these primary abilities will always be less than the number of tests if the latter involve heterogeneous tasks. Third, all of the primary abilities are not necessarily involved in each task (Brody & Brody, 1976).

In 1938 Thurstone proposed a multiple group factor solution that allowed him to empirically determine both the type and number of factors in the matrix. He defined a factor by the number and types of tests that correlate with it. Vernon (1961), however, commented that with Thurstone's centroid technique of analysis "it was natural to expect - not a general factor and small subsidiary group factors - but a number of components of more nearly equal variance" (Vernon, 1961, p. 18). As a guideline or criterion to which his analysis was directed, Thurstone developed the notion of simple structure. Finally, Thurstone advocated the rotation of the factor reference frame as a procedure that would increase the psychological meaningfulness of most factorial solutions.

Thurstone's first large scale study was published in 1941. On the basis of his findings seven primary abilities or factors became the foundation of his theory and measurements. These abilities can be best understood through examples of the tests they represent. Sternberg (1985) has outlined these seven abilities and provided relevant examples of related tests. These are displayed in Figure 2.0.

Additional studies were carried in 1941(a,b) by Thurstone and his wife,Thelma Thurstone. In one study, T. Thurstone (1941b) stressed that the reason behind their search for primary abilities was to obtain a more educationally informative profile of a child's strengths and weaknesses than one can obtain from a single index. This is the same rationale used by Thorndike, Hagen, and Sattler (1986) for dividing the SBIV into four areas of cognitive ability.

Although Thurstone's results initially appeared to contradict Spearman's, both researchers accepted Cattell's 1940 proposition that a second order factor analysis offered a rapproachment between these two positions. Thus in the end, both theorists recognized the existence of group factors as does the SBIV. However, "Spearman continued to assign primary importance to g and to consider group factors or primary abilities as of lesser importance. For Thurstone, the reverse was true" (Brody & Brody, 1976, p. 17).

Few criticisms have been voiced against Thurstone as his contributions were great and his methods were generally considered elegant. Eysenck (1979), however, criticized Thurstone's sampling techniques as they often led to samples that were relatively more homogeneous than the general population. He stressed that such samples reduce the test inter-correlations thereby lowering the variance on the general factor and leading to the elevation of specific or group factors (also see Gould, 1981).

FIGURE 2.0

Thurstone's Primary Abilities

Ability		Tests
i	Verbal Comprehension	Vocabulary tests (e.g., synonyms, and antonyms) reading comprehension tests
ii	Verbal Fluency	Rapid word production tests
iii	Number	Arithmetic word problems
iv	Spatial Visualization	Mental manipulation of figures, symbols, geometric designs
v	Memory	Recall of words, sentences and paired associates
vi	Reasoning	Analogies or series completions
vii	Perceptual Speed	Rapid recognition of symbols, letter crossings

(Modified from Sternberg, 1985)

In summation, the Spearman and Thurstone debate was over the presence of general vs. specific abilities. It was partially resolved with the development of hierarchical factor theories similar to that proposed by Thorndike, Hagen, and Sattler (1986b). It is to these theories we now turn.

Hierarchical Theories of Intelligence

Cyril Burt's Hierarchical Model of Intelligence. "After Spearman, the strongest influence on British psychologists' ideas about intelligence has been that of Sir Cyril Burt" (Butcher, 1968, p. 46). Burt's (1917) Memorandum on The Distribution and Relations of Educational Abilities provided evidence (that Spearman seemingly ignored) that school subjects involved not only a general factor but verbal, numerical and practical group factors as well. It is also here where Burt proposed his simple summation technique of analysis that was later rediscovered in Thurstone's centroid methodology (Nunnally, 1978). His verbal factor was two-fold factor - one aspect involved more complex processes such as composition, history, geography and science and the other less complex tasks such as word recognition and spelling. The second large group factor was viewed as involving practical skills including use of hands, drawing, writing speed, and neatness. In a later study Burt (1939b) determined that the general factor accounted for approximately 28 per cent of test variance and these group factors accounted for approximately 21 per cent each (Vernon, 1961). The remaining variance, in this and any set of measurements, according to Burt, can be accounted for by specific and error factors. Thus, for Burt, any measurement of human abilities consisted of four factors; general, group, specific and error.

Unlike his early contemporaries, Burt was able to accept the importance of Spearman's g along with more recent findings and combine them into a unitary scheme. The central aspect of this scheme is the recognition of intermediate group factors (Vernon, 1953). Burt stated that "the group-factors so far established are not unlike the more specialized faculties: the best attested to appear to represent verbal, arithmetical or numerical, technical or manual, spatial or observation, aesthetic, and musical abilities: and to these some investigators would add memory, speed, constructive imagination (fluency), and various forms of reproductive imagination (visualization, etc.)" (1939, p. 93). The majority of these group factors appear to have stood the test of time and, more importantly demonstrate a strong relationship to those proposed in the SBIV.

It was in this later study that Burt stated his advanced view of the nature of factor analysis. It is a view shared by this writer. He said:

> ...most factorists do not nowadays regard their factors as separate causal powers lodged in definite material structures. They regard them rather as descriptive terms, providing convenient bases for classification and prediction. Nor do they think of them as simple, unanalyzable and sharply demarcated units. Most of the factors tend to be complex and subdivisible, and tend (at any rate in the case of group factors) to show appreciable overlapping. Thus, they are not single or elementary qualities, but patterns of performance; not ultimate properties of the mind, but rather

> provisional categories, often (like other principles of scientific classification) relative to the problem at hand. (1939, p. 93-94).

In 1949 Burt proposed his five level hierarchical model. At the top of the hierarchy was the human mind. The second level consisted of g and a practical factor. Associations were at the third level. Perception and sensation made up the fourth and fifth levels, respectively (Sternberg, 1977). It was here where Burt stated that after the effect of g was removed the next most important task was to separate verbal from nonverbal abilities. This is a practice carried out by most psychometricians today and more specifically those who utilize the Wechsler scales (Fritzke, 1988; Nunnally 1978).

The hierarchical tree of abilities descending in order of generality is now widely accepted by British psychologists including P.E. Vernon (who will be discussed subsequently). So too are Burt's (1939) views about the equivalences of factor analytic results. He noted that discrepant factorial findings are often related to the unyielding interpretation of different methods of analysis. "Each (theorist) has been tempted to criticize any method yielding results a little different from his own" (Burt, 1939, p. 61). Secondly, he viewed factor analysis as a type of average, and that the nature of factors that are discovered often depends on what is averaged. He concluded, that findings across studies will be more similar if proper sampling of traits and subjects are undertaken. Gould (1981) in his much acclaimed work the *Mismeasure of Man* provides an extensive critique of Burt's work including the allegations of data misrepresentations, reification of factors and assumptions regarding the genetic basis of intelligence.

P.E. Vernon's Hierarchical Model of Intelligence. P.E. Vernon (1961) proposed a hierarchical model of intelligence that has been viewed as a theoretical and methodological extension of Burt's work (Gustafsson, 1984; Kail & Pellegrino, 1985). Vernon's model was based on the evaluation of British Service conscripts and the hypothetical integration of previous factorial investigations. His work has received greater attention and acceptance than his predecessor (Nunnally, 1978).

Vernon's hierarchical group factor theory proposed that most daily life performances are attributable to g and to a few highly specific group factors. The role of broader group factors such as those proposed by Thurstone is relatively minor. His views are illustrated schematically in Figure 2.1.

According to this model after the effects of g are removed (by group-factor or by rotation of centroid factors) the positive residual correlations form two major groups - the verbal educational (v:ed) and the spatial-practical-mechanical (k:m) emerge. The v:ed factor can be broken down into factors of a more minor nature such as fluency and divergent abilities (i.e., scholastic abilities). Similarly, the k:m group factor includes perceptual, physical, psychomotor, spatial and mechanical factors that can be further divided by more specific testing. Vernon was, however, critical of multiple factor theorists who he believed carried this process of reduction too far. Citing the views of Humphreys (1962) and McNemar (1964) he notes that group factors are infinitely subdivisible but that the range of performances they account for soon becomes meaningless.

Vernon's views on intelligence testing are well articulated. Importantly, he defends the use of intelligence tests as measures of *g* and that their high v:ed content contributes to their predictive value for many educational and occupational purposes. He states, however, that spatial or mechanical tests are beginning to be distinguished from these *g* tests. He believes that in many intelligence tests group factors are seldom identified in a sufficient manner. He stresses that such tests as the Stanford-Binet (LM), therefore, do not provide reliable diagnostic indicators of separate verbal, numerical, memory or other abilities. He argues for the use and development of tests that provide better assessment of these major factors.

Vernon's intellectual rigor is clearly demonstrated by his position that his model or theory is not necessarily the only interpretation of the current factorial findings on intelligence. He also emphasized that his model did not account for personality and motivational factors or complex interactions of ability factors despite their undoubted

FIGURE 2.1

Vernon's Hierarchial Model of Intelligence

Has been removed because of the unavailability of copyright permission.

(Sattler, 1982, p. 40)

importance. He did, however, conclude that his model had several advantages over existing theories and, it could, along with other methodological issues, account for the diverse findings in factorial studies of human abilities. Vernon's view was shared by Thorndike, Hagen, and Sattler (1986b) who freely incorporated his ideas into the theoretical model underlying the SBIV.

Cattell's Crystallized and Fluid Intelligences. A third hierarchical model has been proposed by Cattell (1963) and later Horn and Cattell (1966, 1982). Like earlier hierarchical theories Cattell's theory has been viewed as an attempt to reconcile Spearman's and Thurstone's factorial theories. However, unlike the previous theories a relatively more central issue for Cattell was the separation of the effects of heredity and environment.

Cattell, like Vernon, accepts the notion of a hierarchy of abilities and the presence of second order factors. Unlike Vernon and British factorists in general, Cattell starts with what he sees as primary abilities and derives second order relationships from them (i.e., bottom up approach). Although, he agrees with other factorists that his process can be continued in a successive fashion he (and subsequently Horn) chose to stop factoring at the second level (Kail & Pellegrino, 1985).

Horn and Cattell (1966) have identified five second-order or "general" factors. The most important two are fluid intelligence (Gf) and crystallized intelligence (Gc). Both intelligences are said to involve perception, comprehension, and abstraction. Gf, however, is said to be involved in unfamiliar tasks and hence is more biologically driven. Gc is said to be primarily involved in tasks that are verbal or conceptual nature and are more experientially driven. The other three second order abilities identified were General visualization (Gv), General fluency (F or Gr), and General speediness (Gs).

Cattell and Horn's hierarchical theory is similar to Vernon's theory (Kail & Pellegrino, 1985). While claiming that their model is superior Horn and Cattell (1968) indicate that Vernon's v:ed should correspond to their Gc and that k:m corresponds to a mixture of Gf and Gv (also see Gustafsson, 1984). According to Gustafsson (1984) "if a third level representing the g factor is added to the Cattell-Horn model the most essential point of difference between the two major hierarchical models would be resolved" (p. 184). Indeed, Cattell's (1987) most recent work suggests the presence of such a factor.

In 1987 Cattell reviewed his theory in an attempt to reconcile it with current psychometric research. It is here where he proposed his triadic theory of abilities. The theory has identified three levels of abilities. These have been summarized below:

1. capacities (gs) general powers operating through all brain action to effect all cognitive performances. In addition to gr the ability to grasp relationships and correlates, there are at least two other gs - speed, and gf - fluency or retrieval capacity.
2. provincials (ps) powers having to do with the functioning of particular sensory or motor systems
3. agencies (as) aids or acquired cognitive skills based on education and training, and proficiencies based on particular interests

(Cattell, 1987: 408).

Cattell's (1987) most recent work also attempts to clarify his position that the triadic theory is not simply "a statement about factor strata levels in a hierarchy" (p. 367). He admits, though, that most of his primary abilities are agencies and most general higher order factors are capacities. Crystallized intelligence is identified as a capacity. He stresses that in reality *g* is simply a mathematical concept. Although he chooses not to discuss heredity vs. environmental origins, he states that GF is largely hereditary and particular patterns of the agencies grow from it. Cattell is, however, less clear at what level the provincials should appear at (Fritzke, 1988).

Undheim (1976, 1981a,b) has conducted a number of factor analytic studies to test the propositions of Cattell and Horn. These studies support the broad distinction of fluid and crystallized abilities and their correspondence to Vernon's group factors. Support was not found for the elevation of GF to the level of Spearman's g, and the developmental propositions put forth by Cattell and Horn (also see Guilford, 1967; Humphreys, 1967; and Vernon, 1961).

J. P. Guilford's Structure of Intellect (SOI) Model

A multifactor theory that is not directly related to the SBIV but that has achieved much prominence is the Structure of Intellect Model proposed by J.P. Guilford (1967, 1982; Guilford & Hoepfner, 1971). This theory rejects the concept of general intelligence and the belief that intelligence can be represented by a small number of group factors arranged in hierarchical form. It accepts the notion of primary abilities such as Thurstone's and these are incorporated into Guilford's model. However, they are divided into smaller abilities and added to new ones so that the total number of abilities reaches 150. Like Thurstone, Guilford originally hypothesized that his abilities were uncorrelated or independent (Sternberg, 1977).

Guilford (1967) also proposed that every mental task consists of three parts; an operation, content and product. Five kinds of operations: cognition, memory, divergent production, convergent production and evaluation are described. These operations are performed on five kinds of content; visual, auditory, symbolic, semantic, and behavioral. They result in six possible products; units, classes, relations, systems, transformations, and implications. as the subcategories are each defined independently 150 (5x5x6=150) mental abilities are proposed (Sternberg, 1985). Figure 2.2 is a graphic representation of this model and the codes used by Guilford.

FIGURE 2.2

Guilford's Structure of Intellect Model

Has been removed because of the unavailability of copyright permission.

(Sattler, 1982, p. 32)

Guilford (1985) states that although the cells are independent (i.e., each ability is defined by its unique position on each of the three dimensions) they can be psychologically correlated. Guilford also proposed that this structure makes it possible, through the use of factor analytic techniques, to develop tests that "load substantially on only one of the many different factors that can be extracted in a single study" (Brody & Brody, 1976, p. 40). By the most recent publication Guilford claims to have demonstrated the existence of 105 of his 150 possible factors (Sternberg, 1985).

Guilford's theory has received sharp criticism. First, the theory as originally proposed required the radical rejection of all major notions of intelligence. Secondly, the statistical methods employed to verify his theory have been questioned (Kail & Pellegrino, 1985). The theory has also been difficult to verify and the few attempts to do so have found questionable results (Brody & Brody, 1976). Fourth, its practical implications for the field of education are limited by the sheer number of abilities proposed (Undheim & Horn, 1977).

It is, however, important to note that Guilford has recently made some concessions and has accepted the view that higher order factors need to be considered. Thus, Guilford himself is moving towards a hierarchical model of human abilities. Finally, Guilford's work has contributed much to research and debate in the area of human abilities particularly "by focusing attention on the issue of mental 'processes' and 'products' as important topics in understanding task performance and human abilities" (Kail & Pellegrino, 1985, p. 41).

An attempt to reconcile 80 years of Factorial work on human intelligence: The Hierarchical Lisrel Based (Hili) Model

Gustafsson (1984) has attempted to reconcile the differences between the above noted factor analytic models. He views the main differences as being largely derived from the method of analysis and choice of rotation. The Lisrel technique, a type of confirmatory factor analysis, was applied to the results of a test battery consisting of 16 tests that were administered to 1000 sixth grade students to test this view.

Gustafsson (1984) concluded that the results of his study provided excellent support for "primary factors as in the Thurstone tradition as well as for the second-order factors . . . Crystallized intelligence, and General visualization hypothesized by Horn Cattell" (p. 179). Support was also obtained for the second-order factor of Fluid intelligence which Gustafsson sees as being identical to a third order g factor.

On the basis of these results Gustafsson proposes a three level hierarchical theory. The highest level is occupied by a general factor similar to that proposed by Burt (1955), Spearman (1904, 1927), and Vernon (1961). Two broad factors reflecting Verbal and Figural information abilities (i.e., similar to Cattell and Horn's 1966 Gf, Gc and Gv and Vernon's 1960 v:ed and k:m) are described. Thurstone's primary abilities are found at the third level. Thus, Guilford's early SOI model is the only major factor theory that fails to be accommodated within the Hili model. Gustafsson is, however, hopeful that Guilford's recent acceptance of higher order factors may lead to a reformulated SOI model.

Gustaffsson's (1984) Hili model provides preliminary theoretical and empirical support for much of the theoretical model underlying the SBIV. His recognition of the need for broad sampling of diverse abilities as well as the advantages of confirmatory analysis is admirable. The Hili model appears to be a logical outgrowth of over eighty years of factor analytic research in the area of human intelligence.

Emerging Theories

Information Processing Theories. Since the 1970's an increasing number of ability theorists have accepted an information processing or cognitive view of human intelligence (Sternberg, 1979). Like differential psychologists, information-processing theorists approach intelligence in various ways and at various levels. Information processing theorists, however, typically attempt to address the following five questions:

> 1. What are the mental processes that constitute intelligence performances on the various tasks?
> 2. How rapidly and accurately are these processes performed?
> 3. Into what strategies for task performances do these mental processes combine?
> 4. Upon what forms of mental representation do these processes and strategies act?
> 5. What is the knowledge base that organize, into these frames of representation, and how is it affected by, the processes, strategies, and representations that individuals use?"
>
> (Sternberg, 1985, p. 1)

Information-processing psychologists, therefore, share a different objective than psychometricians. The former emphasize dynamic processes and the latter static structures. Information processing psychologists study task differences. Psychometricians are concerned primarily with individual differences (Sternberg, 1985).

Butterfield (1986) has reviewed the literature on cognitive differences among people. He has identified four major findings relating to level of development and intelligence. He found that "younger and less intelligent individuals; a) have less elaborately organized knowledge bases, b) use fewer simpler, and more passive processing strategies, c) have less metacognitive understanding of their own cognitive systems and of how the functioning of these systems depends upon the environment, and d) use less complete and flexible executive processes for controlling their thinking" (p. 46).

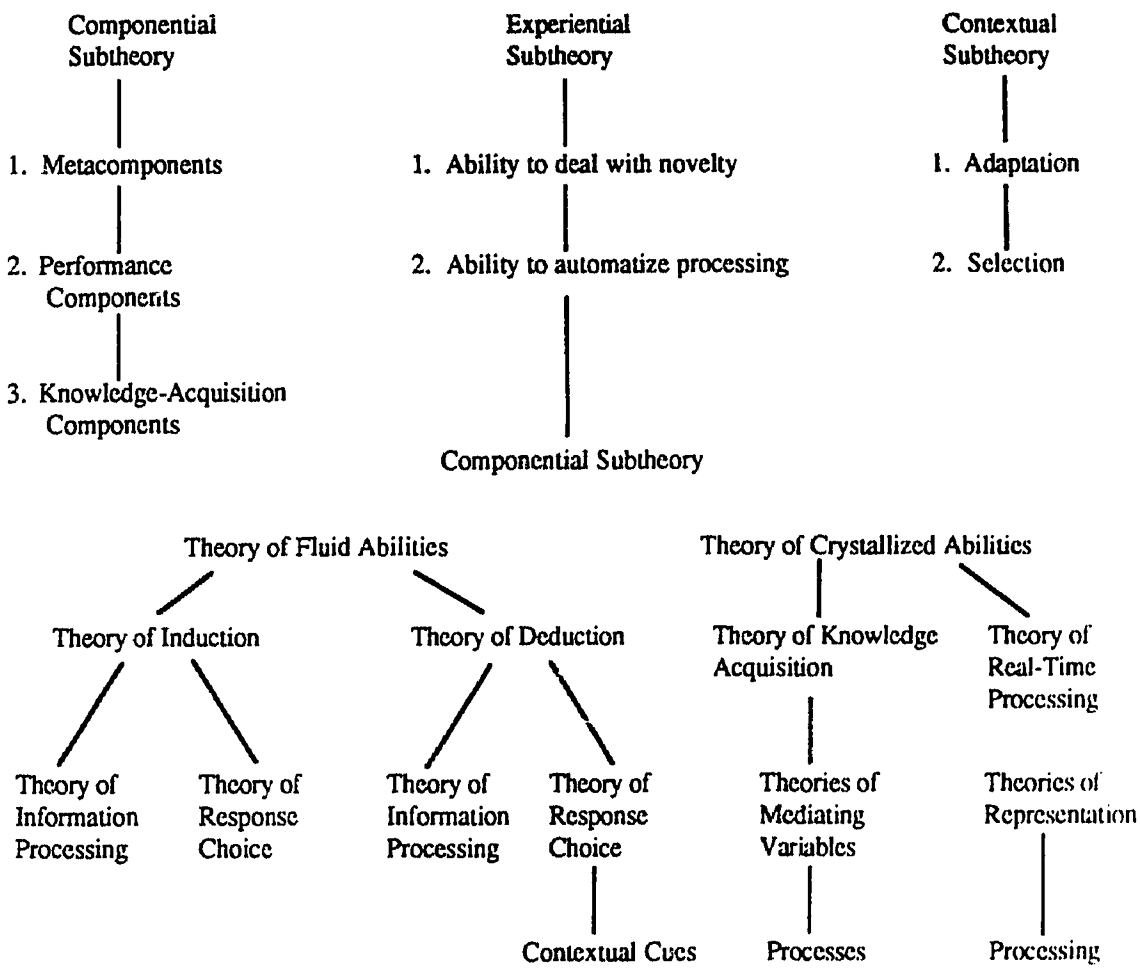
Sternberg (1977) proposed an information processing theory called Componential Analysis. The theory consisted of four essential parts. These are a) the identification of the type of components involved in task performance, b) specification of the combination rule for these components, c) determination of the order of component execution, and d) the mode of component execution. According to Sternberg, several componential theories or models are required to account for intelligent behaviors. A critique of this theory is provided by Gustafsson (1984).

Sternberg (1985) has recently proposed a triarchic theory of intelligence which incorporates his componential theory of intelligence as a sub-theory. Two more subtheories, the contextual and experiential, are added. The theory is an attempt "to specify the loci of human intelligence and how these loci operate in generating intelligent behavior" (p. 317). The theory is presented diagramatically in Figure 2.3 and its incorporation of a diverse range of knowledge on human abilities is readily apparent.

In summation, Sternberg (1985) states that the most widely respected measures of human intelligence including the Stanford-Binet (LM) fail to accurately measure even their creator's conceptions of intelligence. He states that an adequate assessment tool should measure at a minimum each of his three subtheories of intelligence. Sternberg (1985), therefore, stresses that current psychometric tests are restricted to measuring outcomes of knowledge acquisition and present levels of performance functioning. They fail in his mind to address; (a) real world adaptation, (b) novel tasks and situational problem solving, and (c) metacompential planning and decision making. Sternberg (1985), however, fails to outline a concise, efficient, and effective device to fill this assessment role.

FIGURE 2.3

Triarchic Theory of Intelligence



Task Models

Analogies Linear Syllogisms Learning Processing
Series Completions Categorical Syllogisms Vocabulary Vocabulary
Classifications Conditional Syllogisms
Metaphors
Causal Inferences

Gardner's Theory of Multiple Intelligences. Gardner (1985) proposed that there is substantial evidence for the existence of several relatively autonomous human abilities. His findings are based on what he calls a subjective factor analysis of several unrelated groups including "studies of prodigies, gifted individuals, brain-damaged patients, idiots savants, normal children, normal adults, experts in different lines of work and individuals from diverse cultures" (Gardner, 1985, p. 9). His work was guided by what he calls an "optimal definition of intelligence." He states:

> In our view, an Intelligence is an ability or set of abilities that permits an individual to solve problems or fashion products that are of consequence in a particular cultural setting. The problem solving skill permits one to approach a situation, in which a goal is to be obtained, and to locate and pursue appropriate routes to that goal (Gardner, 1985, p. 165).

From this ecological and ethnological perspective seven candidate intelligences are proposed. These are the; linguistic, logical-mathematical, musical, spatial, bodily-kinesthetic, interpersonal and intrapersonal. Although this list is not viewed as being definitive, Gardner sees these seven as providing a reasonably accurate account of intelligences that are valued and/or required to solve problems or fashion products.

Gardner's theory has been criticized on several grounds (Walters & Gardner, 1986). First, several of Gardner's intelligencies are viewed more commonly as 'gifts.' Secondly, his data collection techniques are less than rigorous and he appears to choose only those facts that support his position. Third, his theory is nonfalsifiable. Fourth, the materials used to evaluate each domain are often very different and logistically difficult to implement in systematic fashion. The processes underlying these domains are left uncertain and Gardner cannot adequately explain their general correlations (Walters & Gardner, 1986). In light of these and other criticisms it is somewhat surprising that Gardner's work has generated such interest and acclaim.

Factorial Foundations of the Stanford-Binet (Fourth Edition)

The early Binet-Simon scales were not based on a well formulated theory of intelligence. Their purpose was solely to delineate or quantify levels of intelligence. Once this goal was achieved the focus on general intelligence became more prominent. It was not until the American revisions became popular that the search for separate factors such as memory and sensory discrimination took hold (Terman & Merrill, 1973). According to several subsequent statistical investigations, the Stanford-Binet test data or view of intelligence can be portrayed in terms of general, group, and specific factors. The purpose of this section is to simply report these factorial findings to determine if there is empirical support for the hypothetical factor structure proposed by the SBIV. If the factor structure of the SBIV is to have diagnostic value a certain amount of comparability between studies is necessary (Thompson, 1984). It will, however, be evident that not all studies load on the same group factors. Conclusions of any factor study are limited by the method of factor analysis selected (including factor extraction criteria and rotation procedure), the nature of the sample and the range/type of test items. The latter is particularly problematic for the Stanford-Binet as items within specific subtests are not factorially equivalent at every age (Jones, 1949; McNemar, 1942).

Thompson (1984) has provided a summary table of the factor studies done prior to the 1986 SBIV. The results of the factorization of the SBIV will be summarized in a subsequent section. For ease of interpretation the results of Thompson's study have been included in a revised table. As can be seen from Table 2.0 both consistencies and inconsistencies between studies are present.

It is evident from Table 2.0 that a strong general factor is typically found (Burt & John, 1942ab; Hallahan, Ball, & Payne, 1973). Several group factors are frequently identified including; Verbal, (Burt & John, 1942a,b; Hallahan, Ball, & Payne, 1973; Jones, 1949, 1954; McNemar, 1942; Ramsey & Vane, 1970; Thompson, 1984; Wright, 1939), Nonverbal-visual spatial (Burt & John, 1942; Hallahan, Ball, & Payne, 1973; Jones, 1949, 1954; McNemar, 1942; Ramsey & Vane, 1970; Thompson, 1984; Wright, 1939), Number, (Burt & John, 1942; Jones, 1949; McNemar, 1942; & Wright, 1939), and Memory (Burt & John, 1942; Jones, 1949, 1954; McNemar, 1942; Ramsey & Vane,

TABLE 2.0

Summary of Previous Factor Results on Binet Scales Prior to the Fourth Edition

Study	Age of Subjects	Factors
Burt & John (1942)	10-0 to 11-6	G, Maturity, Verbal, Vocabulary, Number Comprehension, Spatial, Memory
Wright (1939)	10-0 to 10-11	G, or Maturity, Verbal, Number, Spatial,Reasoning Induction
McNemar (1942)	2-1 to 18-11	G only - some indication of motor, naming, numerical, verbal, memory, and problem tests forming distinct groupings
Jones (1949)	7-0 to 7-11	Verbal, Reasoning, Memory, Number
	9-0 to 9-11	Verbal, Spatial, Reasoning, Memory
	11-0 to 11-11	Verbal, Spatial, Memory
	13-0 to 13-11	Verbal Reasoning I, Reasoning 2, Memory, Visualization, Spatial
Jones (1954)	13-0 to 13-11	Verbal I, Verbal 2, Verbal III, Rote Memory, Meaningful Memory, Spatial Reasoning, Closure, Perceptual Carefulness
Ramsey & Van (1970)	3-0 to 7-11	Verbal, Visual and Visual-Motor, Visual Ability and Judgement, Acquired Knowledge, Control of Impulsivity, Visual Imagery
Hallahan et al. (1973)	3-0 to 6-0	G, Visual Ability and Judgment, Verbal
Thompson (1984)	3-12	Verbal Reasoning, Nonverbal Reasoning Visualization, Memory

1970; and Thompson, 1984.) Thus, it can be seen that the proposed factor structure of the SBIV receives some preliminary support from previous factorial studies.

Construction of the Stanford-Binet (Fourth Edition)

The above review suggests that the SBIV has considerable historical, theoretical and empirical support. Indeed, this support led Thorndike, Hagen and Sattier (1986) to select the four areas of cognitive ability that are to be appraised; Verbal reasoning, Quantitative reasoning, Abstract/visual reasoning and Short term memory. This evidence also led to the decision to continue to provide a composite or overall score. A major revision was the acceptance of an item structure much like the Wechsler scales where the items are grouped into subtests.

The authors maintain that the two major strengths of previous Stanford-Binet editions are incorporated into the new test. One of these is the adaptive-testing format that allows each individual to be tested in a range best suited for their level of ability. The second major strength is that the SBIV remains a continuous scale for appraising cognitive development from age two to adult.

An effort was also made to maintain continuity with previous Stanford Binet tests by retaining as many earlier items as possible. Four criteria for item selection were employed. Retained items had to be (a) an acceptable measure of verbal, quantitative, or abstract visual reasoning or short term memory, (b) scorable in a reliable fashion, (c) relatively free of ethnic or gender bias, and (d) function adequately over a wide range of age groups. Items that met these criteria were then subjected to empirical evaluation. The addition of new items and the revised format thus led to changes in content, administration and scoring procedures.

Evaluation of the Stanford-Binet (Fourth Edition)

Osberg (1986) notes the the SBIV was based on the largest standardization sample of any individualized test to date. The total sample size was 5,013. An attempt was made to create a stratified sample comparable to that of the 1980 U.S. census. Primary stratification took place at the school level where children were randomly selected. The five standardization variables were geographic region, community size, ethnic group, age, and gender. Despite these procedures, children whose parents occupied managerial/professional positions and who were college educated were over-represented. A weighting procedure was used in an attempt to correct these deficiencies in the standardization sample data. Sandoval and Irvin (1988) have criticized both the rationale behind and the accuracy of the weighting procedure employed. They believe that any procedure is "a weak justification for a less than desirable methodology in the standardization" (p. 159). Their critique is overly harsh as the procedure used to correct standardization weaknesses is considered common psychometric practice.

The literature on the SBIV is continuing to grow. It can be grouped into three areas. These are practical considerations, technical adequacy and empirical validation. Sandoval and Irvin (1988) have summarized many of SBIV practical concerns.

> 1. Although the entry-level process is relatively simple once understood, many of the subtests have both entry-level directions and continuing directions because different examinees will start each subtests from different levels. In the process of going in reverse order through the items in establishing the basal for a given examinee, the examiner may wonder if entry-level instructions for earlier items on the test should be given to the examinee. The examiner is left without guidance on this matter.
> 2. Instructions and scoring guidelines in the item books are sometimes inadequate, ambiguous, or misplaced. Subtests that require particular examiner attentiveness are Copying, Memory for Objects, and Pattern Analysis....
> 3. Clerical errors can easily result from the use of the tables in the manual that convert subtest SASs to area SASs and area SASs to the composite SAS. There is probably no better way to display these scores, but the user needs to be exceedingly careful lest a wrong column or line by used for the particular conversion needed.
> 4. The protocol form itself is cumbersome (40 full-size pages).
> 5. The quality control of the materials in the kit has been low.
>
> (p. 58 also see Barnard, 1987; Sim, 1987; Slate, 1986; Telzrow, 1987; and Vernon, 1987).

Shortly after the release of the SBIV, the Executive Board of the National Association of School Psychologists (NASP) passed a resolution that it not be used. The resolution was based on the failure of the publishers to distribute technical manuals and appropriate correction procedures to the errors present in the original manuals (Fagan, 1986a, p. 1; also see Barnard, 1987; Slate, 1986). NASP rescinded this resolution in September 1986 when their concerns were rectified (Fagan, 1986b).

Sandoval and Irvin (1988) and Vernon (1987) have reviewed the reliability data provided by Thorndike, Hagen, and Sattler (1986). They conclude that the SBIV has good internal consistency reliability. Almost h. .f of Kuder-Richardson Formula 20 coefficients are in the low to mid-nineties and this held across ages. The least reliable area was Short term memory and, in particular, Memory for Objects. The reliabilities for the Composite Score across all ages was high ranging from .95 (aged 2 years) to .99 (aged 17 years). Naturally, the reliabilities are lower for individual subtests and test-retest reliabilities. Sandoval and Irvin (1988) criticize the test authors for not providing reliabilities and standard errors of measurement (SEM) by ability level. In addition, they suggest that the SEM's proposed should be based on (lower) test-retest reliabilities than on internal consistency reliabilities.

The SBIV provides both evidence for its validity that is both supportive and questionable. The test composite demonstrates excellent concurrent validity with the; Form L-M of the Stanford-Binet (.81), Wechsler Intelligence Scale for Children revised (.83), Wechsler Preschool and Primary Scale of Intelligence (.80), Wechsler Adult Intelligence Scale - Revised (.91), and the Kaufman Assessment Battery for Children (.83). The SBIV has also demonstrated good discriminative validity in the identification of special populations (i.e., gifted, leaning disabled, and mentally handicapped).

The validity of the factorial results, however, have been seriously questioned by Osberg (1986), Sandoval and Irvin (1988), Slate (1986), and Vernon (1987). The data support the presence of g in the composite score. The intercorrelations, however. vary from .29 to .73. "Because the intercorrelations are not equal, Reynolds (1988) questioned the authors" recommended practice of using almost any combination of tests to produce a composite score" (Sandoval & Irvin, 1988, p. 160). The results of the unnamed variant of confirmatory factor analysis do not support the presence of four cognitive ability group factors. Vernon (1987) notes that the total group factor content is only 11 per cent. In addition, many of the tests possess low relationships with other tests in their hypothetical area grouping and on the factor they are designated to load on.

Osberg (1986), Sandoval and Irvin (1988), Slate (1986), and Vernon (1987) conclude that although there is adequate validity for the Composite Score the hierarchical theory is inadequately confirmed. Slate (1986) states that it makes little sense to him why the test publishers continue to market the SBIV based on a theory "when the data suggest only partial support, at best" (p. 3). Osberg (1986) notes that now, even the publishers appear to be suggesting that the hierarchical model was proposed simply to guide test development (Sandoval & Irvin, 1988). Vernon (1987) has recommended that more research is needed by Canadian psychologists to further validate it structure.

A handful of researchers have launched empirical investigations of the SBIV. Carvajal and Weyand (1986) administered both the General Purpose Abbreviated Battery of the SBIV and the Wechsler Intelligence Scale for Children-Revised (WISC-R) to twenty-four third grade children. The General Purpose Abbreviated Battery consists of six subtests. These are; Vocabulary, Bead Memory, Quantitative Memory for Sentences, Pattern Analysis, and Comprehension.

The correlation between the SBIV Composite and WISC-R full scale IQ was significant (r=.78). The correlation between SBIV and WISC-R Vocabulary subtests was also significant (r=.65). This was also true of the two Comprehension subtests (r=.42). However, the relationships between the SBIV Quantitative SAS Score and the WISC-R Arithmetic was nonsignificant (r=.14). This was also true for the relationship between SBIV Pattern Analysis and WISC-R Block Design (r=.18). These lower correlations are explained by the authors as being due to the different administration procedures of these "parallel" subtests.

Carvajal et al. (1987a, 1987b) also investigated the correlations among the SBIV, the Peabody Picture vocabulary Test-Revised (PPVT-R) and their Columbia Mental

Maturity Scale (CMMS) for two groups of school children. The first group was a class of 23 third grade children. In addition to the above noted three tests this group was also administered the Goodenough-Harris Drawing Test (GHDT). For this group the correlations between the SBIV and the PPVT-R was statistically significant ($p<.01$). This was also true for the correlation between the SBIV and the CMMS ($p<.05$). The correlation between the SBIV and the GHDT was not statistically significant (Carvajal et al., 1987a).

Carvajal's et al (1987b) second sample was comprised of 21 kindergarten children. Each of the children were administered the SBIV, PPVT-R and CMMS. The correlation between the SBIV and the PPVT-R in this study was .560 and this is similar to the correlation obtained for the third grade class ($r=.601$). The correlation between the SBIV and CMMS was .400 and this is also lower than the correlation obtained for the older group (.477). Importantly, for this younger group the obtained correlation is not statistically significant. On the basis of these two studies Carvajal et al. (1987a, 1987b) conclude that the PPVT-R is likely the test of choice when a screening instrument is required. They also suggest that administration of the CMMS should be restricted to situations where the child's language deficits preclude the use of the SBIV. The use of the GHDT as a screening instrument is not recommended.

Rothlisberg (1987) examined the concurrent validity of the SBIV and the WISC-R. Her sample consisted of 32 nonexceptional elementary school children. Significant correlations were obtained between the SBIV Verbal Reasoning Area and the WISC-R Verbal IQ ($r=.70$), the Quantitative Reasoning and WISC-R Performance IQ ($r=.40$) and the SBIV Composite and WISC-R Full Scale IQ ($r=.77$). Additional significant and nonsignificant positive correlations were found. She concludes that intercorrelations generally supported Thorndike, Hagen and Sattler's (1986b) predictions about the relationship between these two tests.

Lukens (1987) administered the SBIV to 31 mentally handicapped adolescents who had previously been tested by the Stanford-Binet, L-M. The mean interval between testing periods was 17.3 months. The correlation between the two tests was significant (r=.86). Lukens concludes that the compatibility of the two scores supports the use of the SBIV with mentally handicapped populations.

Meloff (1987) obtained a sample of 153 individuals from a Canadian clinic population who had been administered the SBIV. This sample was divided into four groups: learning disabled, mentally retarded, average, and gifted. The results of her study confirmed that the SBIV was able to distinguish learning disabled from mentally populations and gifted from normal populations.

Meloff (1987) also compared her normal achieving sample to the American standardization sample. Significant differences in favor of the Canadian sample were reported for the two samples Verbal Reasoning and the Composite IQ scores. There were also significant differences for eight of the eleven subtests administered. This is interesting as Vernon (1987) suggested that Canadian samples should score higher than their American counterparts because of the large minority group in the U.S. standardization sample (approximately 25%). However, given the 21% of Meloff's sample were non-white, additional factors such as educational standards may be at work.

Clark (1988) investigated the discriminant and convergent validity of the SBIV. The SBIV along with the WISC-R, Wide Range Achievement Test Revised (WRAT-R), Developmental Test of Visual Motor Integration (VMI), and Perception of Ability Scale for Students (PASS) were administered to a clinic sample of 326 individuals ranging in age 2 to 27 years. The sample was divided into four groups; learning disabled, mentally handicapped, normally achieving, gifted preschool.

Clark's (1988) results are significant in several areas. They support Thorndike, Hagen, and Sattler's (1986b) position that the SBIV is able to discriminate between special populations and normally achieving groups. This was also true of the subtest and Area scores of both the larger and General Purpose Abbreviated Battery. Highly significant correlations between the SBIV Composite, Area and subtest scores and the WISC-R IQ and subtest scores were obtained. Correlations between the SBIV and the WRAT were significant. More moderate correlations were found between the SBIV and the VMI. Correlations between the SBIV and the PASS were low. Overall Clarke's results suggest that the SBIV possesses good concurrent, convergent and discriminant validity.

Hofman (1988) examined the comparability of the unabbreviated SBIV and four abbreviated versions (Quick Screen General Purpose, Potentially Gifted, Learning Difficulties). Data was gathered from a sample of 108 children. All correlations were highly significant. Hofman concluded that the abbreviated batteries are excellent estimates of unabbreviated SBIV Composite Scores.

Fritzke (1988) examined the correlation and factorial results of a clinic sample of 168 children who were administered both the SBIV and the WISC-R. His results like Carvajal and Weyand (1986), Clarke (1988), and Meloff (1987) suggests that the SBIV Composite scale is highly correlated with the WISC-R Full Scale IQ. Fritzke's factorial results are divided into three age groups designed to take into consideration the age restrictions of specific subtests.

A two component solution at the 6-0 to 8-11 year level was suggested by the eigenvalue one and scree tests when a quartimax solution was utilized. Fritzke (1988), however, interprets three in an attempt to find a more psychological meaningful result. Fritzke's g factor accounted for a good deal of the total subtest variance. All subtests with the exception of Bead Memory had substantial loadings on this factor. The second factor was identified as Verbal ability. Vocabulary, Comprehension, and Memory for Sentences had moderate loadings on this factor. The third and more tentative factor was identified as a quasi Abstract/visual factor. It was identified primarily by a single subtest (Bead Memory) and its designation as a factor is questionable. As a general rule three variables are needed to identify a factor. Otherwise the loadings may simply be error.

In a varimax solution of the same data two components were identified by Fritzke. The first was a general factor with strong loadings on all subtests except Bead Memory. The second factor was identified as the ability to perceive and organize material with loadings from Bead Memory, Comprehension, Memory for Sentences and Pattern Analysis. Fritzke also identifies a three component varimax solution that appears to measure Fluid Intelligence, Verbal ability and the ability to Perceive and organize material.

For the 9-0 to 12-11 age group the quartimax solution identifies a strong general component. An Abstract /visual reasoning ability receives some support. However, both Bead Memory and Memory for Objects appear to be related to this factor.

The various rotations for this middle group are somewhat more promising in terms of their support fort the proposed SBIV theoretical model. Fluid and Crystallized Intelligence emerged as did a Short term memory factor. However, as is expected with this type of solution a general factor did not emerge and such a factor plays a prominent role in Thorndike, Hagen, and Sattler's (1986) model.

Fritzke's third and oldest group (13 years +) was extremely heterogeneous. Therefore, he advises that his results are very tentative. Both the scree and eigenvalue one criteria suggested a single component, general factor solution. Both rotational solutions were supportive of a single factor solution.

Four main conclusions can be derived from Fritzke's analysis. First, the SBIV appears to be a good measure of general intelligence. Secondly, the four cognitive ability areas defined by Thorndike, Hagen, and Sattler (1986) are not substantiated by Fritzke's methods. However, Fritzke's analysis was based on exploratory rotations. It can be argued that confirmatory analysis would have been more appropriate given the presence of and desire to test a specific theory. Third, Bead Memory appears to measure a Visual/spatial as opposed to Short term memory factor. Finally, additional research using various methodological techniques such as hierarchial analysis is advocated to assist in the further determination and specification of the factor analytic structure of the SBIV.

Sattler (1988) conducted a principal components analysis with a varimax rotation on the SBIV standardization data. Unlike the SBIV Technical Manual and proposed theoretical structure Sattler gives specific recognition to the dynamic nature of the SBIV factor structure. He writes that "because subtests in the ... (SBIV) ... are not continuous throughout the scale and because different subtests are administered at different ages, the factor structure of the scale differs at different ages" (p. 255). It is also important to note that the factor structure may also be different within scales as the nature of the test stimuli may change as the child moves between test levels (e.g., Vocabulary, Pattern Analysis).

Sattler (1988) concludes that his results are supportive of a two factor interpretation at the 2 to 6 years ages. These are Verbal Comprehension and Nonverbal Reasoning/Visualization. Noteworthy is that Bead Memory is assigned to this latter factor. Three factors are reported to emerge at the seven year level. This third factor is identified as a Memory component. These results are outlined along with their corresponding subtests in Table II-8 of Sattler (1988, p. 259). Overall, it is noteworthy that one of SBIV test authors now advocates an interpretative model that is different from that initially proposed and outlined in the 'Technical Manual' (1988, pp. 278-279). Similar to Fritzke's (1988) analysis one can argue that given the SBIV theoretical structure confirmatory factor analytic techniques should have been employed by Sattler (1988). However, Sattler (1988) provides no explanation for his choice of exploratory confirmatory factor analytic techniques.

Keith et al. (1988) have addressed one of the possible criticisms of Fritzke's (1988) and Sattler's (1988) studies. The former have investigated the construct validity of the SBIV using confirmatory as opposed to exploratory factor analysis. Their rationale for selecting this route is well articulated and is based on the belief that confirmatory factor analysis "provides a much stronger test of the underlying structure of a test instrument than does exploratory factor analysis and is especially useful for tests such as the new Binet that are based on an explicit underlying theory" (Keith, et al., 1988, p. 258).

The purpose of Keith's et al. study was to determine the extent to which the four cognitive areas of ability are found. They argued that:

> one can conclude little from the factor analytic results contained in the ... (SBIV Technical Manual).... The factor loadings appear to be much smaller than would be expected, but the the vague reference to the procedure used makes it difficult to know exactly ... 'd be expected. Without more detail as to how this "variant of confirma... ...lysis" (p. 52) was conducted, the findings reported cannot be inte... re garded as unequivocal support for the test's factor structure (p. 2...

To achieve grea... clarification on the structure of the SBIV, first order, confirmatory factor analyses were performed using the LISREL VI computer program. Models were specified in accordance with the hypothesized four areas of cognitive ability and their respective subtest groupings. The four cognitive areas were allowed to correlate with each other as this is consistent with the theory underlying the SBIV and reflects the presence of *g* or general intelligence. Four age groups (2-23, 2-6, 7-11, 12-23 years) were analyzed. Following this less structured versions of the model were analyzed provided that each of these models held true to the constructs presumably measured by SBIV.

The results of Keith et al.'s goodness-of-fit analysis on the entire SBIV standardization sample for the first two levels suggest an adequate fit between the model and factorial findings. There were, however, some inconsistences with the SBIV theory. Verbal and Quantitative reasoning Areas were hypothesized by Thorndike, Hagen and Sattler (1986) to require crystallized abilities. Therefore, these two reasoning areas were predicted to correlate more highly than any combination of crystalized and fluid reasoning scales. For the entire standardization sample "the correlation between verbal and quantitative was the lowest factor intercorrelation (.79), however, and the quantitative factor correlated most highly (.92) with the abstract/visual factor, which required fluid abilities according to the Binet theory" (1988, p. 261). Despite these inconsistencies the hypothesized intercorrelations are quite high and certainly support the presence of a *g* factor (almost all first-order factor correlations were above .75 and several were above.90) underlying all scales and tests.

A second analysis was conducted that tested the goodness of fit of a more relaxed theoretical model. This model allowed for the Memory for Sentences, Bead Memory and Absurdities tests to load outside their hypothesized Reasoning Areas. A significantly better fit was obtained under these parameters. Importantly, Keith et al. (1988) argued that the relaxed model parameters "were consistent with the constructs that the Binet presumably measures" (p. 259). The authors have summarized the revised model findings:

> Bead Memory loaded highest on the abstract/visual reasoning factor (.49) rather than on short-term memory (.27). Memory for Sentences loaded almost as highly on the verbal reasoning factor (.40) as on the short-term memory factor (.43). Absurdities continued to load more highly on the verbal factor (.47), with a lower loading on abstract/visual (.27). The relaxation of the model slightly lowered the correlations among several of the factors, but the correlation between the abstract/visual and quantitative factors was still .90 (p. 264).

Importantly, the goodness of fit for the second model was found to be significantly better than the strict model proposed by Thorndike, Hagen and Sattler (1986) for the standardization sample.

Confirmatory analysis was then conducted on each of the three age groups. For the 12-through 23 year old age group the results were substantially similar to those for the entire standardization sample. The strict model correlations among the first order latent factors were high. Similar to the entire standardization sample the highest correlation was between the Quantitative and Abstract/visual factors (.93); a somewhat lower correlation (.85) was found between the Verbal and Quantitative factors.

Similar to the entire standardization sample the relaxed model provided a significantly better fit for the 12 to 23 age group. When specific subtests were allowed to load on more than one factor, "Bead Memory loaded more highly (.52) on the Abstract/visual than on the Memory (.25) factor. Absurdities loaded almost equally on the Verbal (.37) and the Abstract/visual (.40) factors, and Memory for Sentences loaded only slightly higher on the Verbal (.44) than on the Memory (.38) factor. As in the strict Binet model, the correlation between the Quantitative and the Abstract/visual factor was the highest (.92); the Quantitative - Verbal correlation was somewhat lower (.83)" (p. 265-266).

The confirmatory factor analysis results for the 7 to 11 year old age group indicate that, in general, the strict theoretical model fits the standardization data. However, some inconsistencies were found. Specifically a correlation of 1.0 was found between the Quantitative and Abstract/visual factors. This suggests that for elementary school aged children these factors may be identical. This finding is inconsistent with the SBIV theoretical model which hypothesizes that the Quantitative factor measures crystalized intelligence whereas the Abstract/visual factor measures fluid intelligence.

Similar to the entire standardization sample analysis the results of the relaxed model analysis for the 7-11 year old age group indicated a significantly better fit, than when the strict model was employed. The relaxed factor model also saw a reduction in the correlation between the Quantitative and Abstract/visual factor to less than 1.0. The obtained value (.97) was still, however, higher than the Verbal-Quantitative correlation (.82).

The analysis for the youngest age group (2-6 years) was constrained by the fact that only one test (Quantitative) comprises the Quantitative Reasoning Area Score. Keith et al. (1988) dealt with this problem by restricting the Quantitative Reasoning - Quantitative test factor loading to the square root of the reliability of the Quantitative test and restricted the unique variance of the Quantitative test to the complement of reliability. There were, however, difficulties in approximating within the constrained SBIV model despite these adjustments. Specifically, a perfect correlation between the Memory factor and the Verbal and Abstract/visual reasoning factors was found. Such a finding, according to Keith et al. (1988) suggests that "short-term memory may be indistinguishable from reasoning for preschool children, at least by the Binet tests" (p. 269). A subsequent analysis of a No-Memory model provided an "excellent fit to the 2-through 6-year old standardization data" (p. 269). All of Factor loadings were significant and high. Noteworthy, was the fact that the memory tests loaded at a somewhat lower level than the other tests. Importantly, and like the other analyses, the correlation between the Verbal and Quantitative factors (.63) was somewhat lower than the Quantitative-Abstract/visual correlation (.80) and the Verbal-Abstract/visual correlation (.75).

The final analysis by Keith et al. (1988) addressed the question raised by Sandoval and Irvin, (1988) and Slate (1986). That is, does the SBIV measure anything of significance other than *g*? The *g* only model constrained the correlations among first order factors to 1.0, thereby assuming that all factors measured the same construct, namely general intelligence. The g only model was found to provide a significantly worse fit to the standardization data than did any of the strict models for the four age groupings. Keith et al. (1988) concluded that although the Binet possessed a strong *g* component their results suggest that other constructs in addition to g are present.

Keith et al. (1988) conclude that their results generally support the four cognitive ability factors, hypothesized to exist four the SBIV. The one exception to this finding is the 2-6 year old age group where it was difficult to isolate a memory factor separate from the reasoning factors. The presence of a strong g component is also supported. The first order factor correlations are less supportive of the second level of the SBIV theoretical model which combines the scales into measures of Crystallized or Fluid abilities. Keith et al. (1988) suggest that caution is needed in interpreting this level and recommend that hierarchical factor analytic research be conducted to further investigate this aspect of the SBIV theory.

An Introductory Note on Factor Analysis

It is evident from the theoretical and empirical review that factor analysis is now defined by a large number of methods and a broad collection of mathematical techniques. The complexity of the procedures are advancing so rapidly that even the regular user has been required to consult specialists from time to time (Nunnally, 1978). This thesis will approached these techniques as conceptual or logical methods (Burt, 1955). The technical or mathematical aspects of these procedures is discussed only when they specifically determine or restrict the interpretation of findings.

Mathematically, factor analysis is a data reduction technique. It is a method that allows for the reduction of many measures to a smaller, more comprehendable number (Keith et al., 1988). It is a condensation or distillation process. Baird (1987) has developed the concept of *exploratory* factor analysis as an information-transforming instrument. "Such instruments transform and summarize information input into useful form more useful for further research in the output" (p. 320). Baird has compared exploratory techniques to microscopes as both serve to make information more accessible to human examination.

Psychologically, factor analysis attempts to serve two fundamental purposes. First it is "designed to identify a minimum number of latent variables or "factors" in a set of data that will most parsimoniously, and from the standpoint of psychological interpretation most intelligibly account for covariation among variables in the data" (Caroll 1985, p. 25). Keith et al. (1988) describes this process as essentially being the determination of discriminant and convergent validity. Those tests that measure something similar form a factor, whereas tests that measure something different form other factors. Factors are labeled according to what the researcher perceives the measures, on a particular factor, share in common.

The second fundamental purpose of factor analysis for psychologists is to determine if these factors form a pattern or structure (Carroll, 1985). That is, the relationships among factors is examined to determine if some are more general than others. A number of ancillary objectives exist (i.e., calculation of factor scores or factorial composition of variables) but they all share the overriding goal of systematic description or classification.

There are two broad types of factor analysis; exploratory and confirmatory. Exploratory analysis is the earliest and perhaps most commonly known type. This method is generally used when the researcher has no or few preconceived notions about the data. Essentially these techniques "let the data set speak for themselves" (Carroll, 1985, p. 26).

Confirmatory factor analysis is a more recent development. It allows for the statistical testing of hypotheses. It is used when researchers possess theoretical models which are often based on previous factorial work. This method of analysis has been used in development of the SBIV and it is to these results we now turn.

SBIV Factorial Methods and Results

Thorndike (1987), in a personal communication, has outlined the techniques used to evaluate the hypothetical structure of the SBIV. He described his procedures as being "slightly idiosyncratic" and "not mathematically elegant." Importantly, no formal attempt was made to evaluate the accuracy of fit between his data and the hypothesized structure of the SBIV. Factor procedures were stopped when the matrix residuals were perceived to be small. Thorndike's description of his procedures is outlined below.

> Basically, the variables were assigned to their pre-specified clusters. Then by a series of iterations the G-factor weights were determined such that the cross-cluster correlations were minimized. That is, for each variable the sum of its correlations with all of the other variables from other clusters was made as nearly zero as possible. Subsequent, factors were fitted to the matrix of residuals, after the correlation due to the G-factor had been removed. This was done one factor at a time, starting in each instance with the Verbal cluster of variables. The residuals among variables that had been designated Verbal were analyzed, and loadings were determined that would best account for those residuals. Once these had been determined a loading was determined for each other variable that would account, (on average) for that variable's residual correlations with the Verbal tests. Usually these loadings were very small, but Memory for Sentences, for example, had a noticeable (and reasonable) verbal loading. This procedure was repeated for Memory, Abstract/Visual and Quantitative, in turn, where any evidence for a cluster made such an analysis seem fruitful (1987, p. 2).

Therefore, Thorndike's analysis thus starts with a target matrix, based on prior analyses utilizing conventional factor analytic procedures. He describes his procedure as an "old-style Thurstone centroid analysis" with the graphical rotation of results to approximate simple structure and positive manifold. The assumption is that the clusters of tests in the target matrix share only the general factor. Such a procedure attempts to make each of the clusters stand on their own.

Thorndike (1987) was able to admit that there were some problems with his factorial results. He stated that the Matrices test appeared to be almost completely comprised of g and that it did not really fit into the Abstract/visual cluster. He also noted that Bead Memory was somewhat of an anomaly as it did not compare well with the other memory tests. He believed that it may call for simultaneous as opposed to sequential processing and this may explain its uniqueness. In addition, Bead Memory also appeared to have a Visual/perceptual component. Finally, Thorndike recognized the trouble he had in obtaining a Quantitative factor. He saw the later result as being largely due to its substantial g component.

It is apparent, then, that Thorndike himself was aware of the disconfirming nature of some of his factorial results. He has, however, minimized these negative findings and seemingly tries to explain them with after the fact rationalizations. Importantly, no attempt has been made to rectify these discrepant results theoretically or empirically even though a subsequent analysis by Sattler (1988) places the original factor interpretations in further doubt. Administratively, both Bead Memory and Matrics have not been reassigned or removed from their respective cognitive areas. Quantitative Reasoning has remained a cognitive area despite its tenuous existence. Consequently, unquestioning clinicians may assign these variables interpretive significance that potentially does not exist.

Support for the factor structure of the SBIV, therefore, can be found in its theoretical framework and the factorial studies of early editions. The factorial results derived directly from the SBIV, however, are not wholly supportive. Thorndike's (1986b) and Sattler's (1988) results strongly suggest that further research is required before the utility of the four cognitive areas of ability can be accepted. The urgency of this task cannot be overstated as the test is now in use and the potential for interpretative errors is great.

Scatter Analysis

As noted earlier, Alfred Binet was primarily concerned with developing a global index of inter-individual differences in human ability. He was, however, one of the earliest psychometricians to recognize that "the various psychological processes arrange themselves in a different configuration within each individual" (Binet & Henri, 1905, p. 411, cited in Frank, 1983, pp. 66-67). His observations on intra-individual differences were never as fully developed as those on inter-individual differences.

Despite this early recognition of the uneven nature of an individual's abilities, a uniform definition or measure of this "scattering" does not exist. Kramer et al. (1987) have provided generally acceptable definitions of both scatter and scatter analysis. They have written, "scatter refers to the extent of variability of subtest scores that an individual manifests on a single administration of the WISC-R or the SBIV. The use of scatter patterns to generate hypotheses about an individual's potential or to develop educational or clinical interventions is referred to as *'scatter analysis'* (p. 37 italics in original).

Harris and Shakow (1937) provided one of the earliest reviews on research and numerical measures of test scatter. An initial finding was that much of the literature was confictual. They noted "(s)ince these results are based upon the use of different tests, different measures of scatter and different kinds of populations, the differences in results are to some extent to be expected" (p. 135). Such a cautionary statement is still valid today.

Harris and Shakow (1937) also reviewed nine measures of scatter. These were later classified into three groups by Lorr & Meister (1944). The latter wrote:

> The simplest type, range of "scatter", consists in counting the year levels over which both success and failures occur (Wells, 1927; Shipley, 1934). A second type emphasizes area and consists in the number of months credit earned above the basal year (Wallin, 1922). The third type takes both into account, and then the most successful of these consists of multiplying failures below and passes above the mental age by the number and year levels separating each test from the mental age (Pressy & Cole, 1919) (p. 983).

This review is, however, summarized by the warning that research up until their time had failed to demonstrate "any clinical use for ... (these)...numerical measures of scatter" (p. 148).

Lorr and Meister (1941) reviewed the concept of Binet test scatter from the perspective of measurement theory. Five potential sources of scatter were identified. These authors emphasized that:

> ... scatter is a consequence of the lack of the perfect correlation between test items resulting from the presence of error and from the low communality and high specificity of the items... Secondly, there is the fact that items are incorrectly allocated in order of difficulty..1. Thirdly, scatter may be due to the lack of discriminatory power of certain items... A fourth source of scatter may be found in the fact that there is an increase in variability with an increase in absolute mean test performance... A fifth possible cause of scatter is the presence of systematic errors in testing due to language handicaps, sensory defects, special training, lack of cooperation and ambiguous scoring or instructions (1941, pp. 303-306).

Lorr and Meister (1941) also speak to the factorial basis of the Binet. It was used as additional support for questioning the practice of deriving clinical hypotheses from subtests levels. These comments are especially relevant to this study. They wrote:

> First, the factorial composition of an item cannot be prejudged accurately. Such judgements are frequently in complete disagreement with factor analysis results... Secondly, an item might be solved through the use of different abilities by different individuals at different age levels. Thirdly, items may show fairly high loadings on more than one factor so that the failure cannot be attributed to the lack of any one ability. Fourthly, such clusters of items have too low a reliability to have any real diagnostic value (1941, p. 308).

More contemporary reviews of scatter analysis have been provided by Sattler (1982) Reynolds and Clark (1985) and Kramer et al. (1987). Sattler (1982) and Reynolds and Clark (1985) recognize the abundance of literature that suggests this practice is of limited value. Sattler (1982) stated "(f)or the most part...scatter analysis has not been shown to be useful as a valid interpretive device on the Stanford-Binet" (p. 139). He goes on, however, to report several related "findings". Despite his cautionary statement that "these findings have been based on limited samples of children and for the most part have not been replicated" (p. 39) their publication appears to give some credence to this practice.

Reynolds and Clark (1985) are more steadfast in their support of scatter or profile analysis. They argue that this procedure has been both wrongly implemented and criticized. Furthermore, that when consideration is given to both the psychometric and clinical properties of each scale they believe that useful remedial information can be obtained. They cite the work of Kaufman (1976, 1979) and Reynolds and Clark (1983) and Sattler (1982) in support of this practice

Kramer et al. (1987) provide an extensive review of the viability of scatter analysis using either the WISC-R or SBIS. They observe that the results of various studies suggest that while some groups appear to routinely exhibit greater or lesser scatter this information has not proven useful for differential diagnoses or education remediation. They conclude that "the research evidence fails to support the ability of SBIS scatter analysis to predict either membership in clinical groups or academic potential" (Kramer et al., 1987, p. 39). According to these authors such a conclusion is not surprising given that the Stanford-Binet of their day was never designed for such purposes.

This latter comment is no longer valid. The four cognitive areas of ability on the SBIV are intended to supplement the diagnostic and remedial information provided by a global score. The revamping of subtests to a Wechsler scale like format is also designed to facilitate such analyses. Therefore, the question of whether or not scatter analysis is a viable clinical and diagnostic procedure is once again, unanswered.

The purpose of this section is not to take sides on the scatter analysis debate. It is simply to recognize that such a debate exists and to provide preliminary frequency information on the amount of scatter found in a moderately large clinic sample. Matarazzo, Daniel , Prifitera and Herman (1988) state that this is an important first step in exploring the diagnostic importance of scatter (also see Grossman, 1983).

Prior to commencing their own study of the Wechsler Adult Intelligence Scale-Revised (WAIS-R) scatter these same authors reviewed several indices of scatter. The range (difference between highest and lowest scaled scores) was selected as the method of choice. The rationale was

a) it is the easiest to calculate

b) it correlates very highly with the standard deviation (which is by definition the most sensitive of the three scatter indices), and

c) it is nearly as reliable as the standard deviation (Matarazzo, Daniel & Prifitera, 1977, p. 7).

Using the range as their index Matarazzo, Daniel, Prifitera and Herman (1988) calculated the scatter across all 11 subtests and between Verbal and Performance scales for each of the 1,880 in the WAIS-R standardization sample. Full scale score scatter ranged from 2 to 16 points. The average scatter ranged from 2 to 16 points. The average scatter for the Full Scale was 6.7 points. The average scaled score scatter for each of the Verbal and Performance scales was 4.7 points. No relationship between age, sex, race or education was found. A significance positive relationship was found between scatter and intelligence level. Normative tables were provided for interpreting WAIS-R scatter. Several cautionary notes are provided regarding the importance of reliable subtest variance, factorial specificly, adequacy and uniformity of subtest floors and ceilings. A final and salient warning is that:

> Scatter is never interpreted in isolation. Rather it is best, interpreted in light of other objective information available about the examinee, such as previous intelligence test scores from school or occupational records, job history, socioeconomic indices and hospital and other diagnostic information (Matarazzo, Daniel & Prifitera, 1988, p. 11).

Matarazzo, Daniel, Prifitera and Herman (1988) found that the correlations between scatter and age, sex as well as race were small ranging from .02 to -.16. Correlations between scatter and education were also modest (.12 to .24). The highest correlation was between the three IQs and scatter. In general, a direct relationship between scatter and IQ level was observed. Overall, the authors conclude that "a difference between an individual's lowest and highest subtest scores of as much as 9 points on the Full Scale, or

7 points within the verbal or Performance Scale, is not at all unusual" (Matarazzo, Daniel, Prifitera & Herman, 1988, p. 10). Therefore, scatter in these ranges is likely to be less informative. Sattler's (1988) results strongly suggest that further research is required before the utility of the four cognitive areas of ability can be accepted. The urgency of this task cannot be overstated as the test is now in use and the potential for interpretation errors is great.

The Rise and Importance of Ethical Considerations in Psychological Testing

The fact that psychological testing can have significant implications for individuals, their families, institutions and society led to the formulation of a number of professional committees during the latter half of this century to guide the development and use of psychological tests. Three prominent committees published four early documents to guide the development and use of tests. This early work has been chronicled by a recent (1985) American Psychological Association (APA) publication:

> The first of these was *Technical Recommendations for Psychological Tests and Diagnostic Techniques*, prepared by a committee of the American Psychological Association (APA) and published by that organization in 1954. The second was *Technical Recommendations for Achievement Tests*, prepared by a committee representing the American Education Research Association (AERA) and the National Council on Measurement used in Education and published by the National Education Association in 1955. The third, which replaced the earlier two was published by the APA in 1966 and prepared by a committee representing APA, AERA, and (NCME) and called *Standards for Educational and Psychological Tests and Manuals*. The fourth, a revision of the third document, which was prepared by the three sponsoring organizations, was published in 1974. (American Psychological Association 1985, p. 1).

The most recent document written by these three organizations was published in 1985. The *Standards for Educational and Psychological Testing* were formulated with the intent of being consistent with the APA's *Standards for Providers of Psychological Services* (1977) and with *Ethical Principles of Psychologists* (1981). These guidelines are binding for professional psychologists and address the development, publication and use of psychological and educational instruments. They provide a basis for evaluating the quality of practices in these three areas (American Psychological Association, 1985).

Although all three sets of standards are pertinent to intelligence testing and this study, the technical standards for evaluating test construction are salient. Of these, those relating to validity are considered the most important (American Psychological Association, 1985). Three types of validity are commonly distinguished; content-related, criterion-related, and construct-related. Ideally, all are used in the evaluation of a test.

The primary consideration of this dissertation is the construct related variety. However, it must be emphasized that the evidence from content and criterion related validity can contribute to the evaluation of constructs. According to the *Standards* "the choice of which one or more approaches to use to gather evidence for interpreting constructs ... will depend on the particular validation problem and the extent to which validation is focussed on construct meaning" (1985, p. 10). The construct validation approach has been selected because the validation problem central to this thesis is; What do the various subtests mean and what is their relationship to each other and the Composite Score? Specifically, are the four cognitive areas empirically accurate and what are the empirical referents of each subtest?

The *Standards* have stressed the importance of theory in this construct evaluation process:

> The construct of interest for a particular test should be embedded in a conceptual framework, no matter how imperfect that framework may be. The conceptual framework specifies the meaning of the construct, distinguishes it from other constructs, and indicates how measures of the construct should relate to other variables (American Psychological Association, 1985, pp. 9-10).

Theory provides the conceptual framework against which the empirical results are evaluated. Therefore, the process of construct validation is typically on-going, starting at the test development stage and continuing "until the pattern of empirical relationship between test scores and other variables clearly indicates the meaning of the test score(s)" (American Psychological Association, 1985, pp. 9-10). Prior to addressing the theoretical framework guiding the construction of the Stanford-Binet (Fourth Edition) it is important to highlight the ethical standards relevant to this dissertation. These are outlined in Table 2.1.

TABLE 2.1

Ethical Standards Relevant to the Construct

Validation of the Stanford-Binet (Fourth Edition)

Standard 1.1 Evidence of validity should be presented for the major types of inferences for which the use of a test is recommended. A rationale should be provided to support the particular mix of evidence presented for intended uses. (Primary)

Comment: Whether the Stanford-Binet four cognitive areas and respective subtest groupings exist is of primary interest. This standard is particularly appropriate in clinical populations as the questions being asked are related to the determination of specific learning issues (e.g., strengths and weaknesses, remediation strategies, school placement). The inference made by Thorndike, Hagen and Sattler (1986) that the Stanford-Binet (Fourth Edition) is useful for these purposes as are interpretations stemming from each of the four cognitive areas at all age levels.

Standard 1.2 If validity for some common interpretation has not been investigated, that fact should be made clear, and potential users should be cautioned about making such interpretations. Statements about validity should refer to the validity of particular interpretations or of particular types of decisions. (Primary)

Comment: The Standards indicate that it is incorrect to speak in global or unqualified terms about the validity of a test. The Standards also specify that a test is not likely valid for all purposes or under all circumstances. Evaluating a clinical population is a step toward the determination if test inferences from the standardization sample generalize to a clinic population.

Standard 1.3 Whenever interpretation of subscores, score differences, or profiles is suggested, the evidence justifying such interpretation should be made explicit .... (Primary).

Comment: See comment for Standard 1.1.

Standard 1.9 ... (W)hen several scores are obtained from a single test, each purporting to measure a distinct construct, the intercorrelations among the scores for one or more samples should be reported. (Secondary)

Comment: The four cognitive ability areas are designed to provide information on each of their respective theoretical constructs. Relationships among test scores should be consistent with the theoretical model underlying the Stanford-Binet (Fourth Edition)

Standard 1.10 Construct-related evidence of validity should demonstrate that the test scores are more closely associated with variables of theoretical interest than they are with variables not included in the theoretical network. (Conditional)

Comment: See comment for Standard 1.9.

Standard 5.2 Test manuals should describe thoroughly the rationale for the test, state the recommended uses of the test, and provide a summary of the support for such uses. Where particular misuses of a test can be reasonably anticipated, the test manual should provide specific cautions against such misuses. (Primary)

Comment: If support for the Stanford-Binet (Fourth Edition) is weak or restricted in any way it should be reported in such a fashion so as to guard against misinterpretations.

Standard 5.7 Promotional material for a test should be accurate. Publishers should avoid using advertising techniques that suggest that a test can accomplish more than is supported by its research base. (Primary)

Comment: If any or all of the subtests fail to comply to the theoretical framework guiding the construction of the Stanford-Binet (Fourth Edition) the test manual and protocol face sheet should be revised and contain appropriate cautions to guard against misinterpretations.

Standard 7.1 Clinicians should not imply that interpretations of test data are based on empirical evidence of validity unless such evidence exists for the interpretations given. (Primary)

Comment: This standard addresses the fact that test results often have a dramatic and long-term effect on people's lives. Evidence for the four cognitive ability areas and respective subtest allocation, particularly for clinical populations is required before any interpretations can be offered, even simply as hypotheses.

Standard 8.11 Test users should not imply that empirical evidence exists for a relationship among particular test results, prescribed educational plans, and desired outcomes unless such evidence is available. (Primary)

Comment: This standard highlights the role often played by test results in educational environments. According to the Standards, "test results in special education are often used to develop specific educational objectives and instructional strategies that are assumed to remediate a student's educational deficits or to enable the student to compensate for them" (American Psychological Association, 1985, p. 54). If evidence does not exist to support the inferences derived from test results users are allowed to make recommendations provided that their tentative nature is stressed. However, the Stanford-Binet (Fourth Edition) offers the four cognitive areas of ability and respective subtests as if they are empirical givens. Therefore, it is important to determine; to what extent they can be corroberated, or if interpretations must be qualified and/or restricted.

Summation

It can be seen that since the time of early Greek philosophers the exact nature and structure of human intelligence has been elusive. The introduction of the Binet scales has, however, provided psychometricians with an effective tool for establishing general levels of intelligence. The development of factor analysis has assisted researchers in their quest to integrate their theoretical views of intelligence and their practical attempts to measure it. A concomitant rise in ethical issues in psychological testing has renewed the desire for theoretical formulations underlying assessment instruments. As we have seen, however, few researchers have been able to fully integrate their views of intelligence with their empirical referents. This is seemingly true for Thorndike, Hagen and Sattler (1986) and the SBIV.

CHAPTER 3

RESEARCH DESIGN AND PROCEDURES

Introduction

The purpose of this chapter is to outline the instrument, participants, procedures, and research questions that are employed. Primary consideration here is given to the nature of the SBIV, sampling adequacy, choice of age groupings, missing data, procedures, determination of what components to retain and study limitations. Finally the specific research questions that are to be addressed are outlined.

Instrument

It is clear from the preceding chapters that the SBIV is substantially different from previous editions. As noted earlier and illustrated in the *Technical Manual* (Thorndike, Hagen, & Sattler, 1986b), the purpose of the SBIV is to provide, in addition to a composite or overall score, information that would be useful for individual educational diagnostics and remediation. The primary focus of earlier editions was on inter as opposed to intra-individual differences. An ancillary purpose of the SBIV is to reduce the emphasis on verbal skills, especially at older age levels of the scale. Utilizing current theory and research the three-level hierarchical model of cognitive abilities, displayed in Chapter 1 (see Figure 1.0), is proposed.

Four areas of cognitive abilities within the three-level hierarchical model are introduced. At the top level is a general reasoning factor. The second highest level is comprised of three broad factors, Crystallized abilities, Fluid-analytic abilities and Short-term memory. The third level is made up of more specific factors Verbal Reasoning, Quantitative Reasoning and Abstract/visual Reasoning.

Like the theory, the structure of the SBIV is considerably different from earlier editions. Although the SBIV continues to assess individuals ranging from 2 years to adulthood it is divided into various subtests, of similar content, much like the Wechsler scales. The SBIV is comprised of a total of fifteen such tests. Each test is assigned to a scale that it is theoretically believed to gauge. The fifteen tests are described below and are grouped according to the scale they are said to be a part.

> Verbal Reasoning (Crystallized Abilities)
>
> *Vocabulary.* Pictures or visually and orally presented words are defined.
>
> *Comprehension.* Initial comprehension items consist of pointing out body parts on a picture, whereas later items require verbal responses to questions (e.g., "Why do people have stores?").
>
> *Absurdities.* Absurd aspects of pictures are described and explained.
>
> *Verbal Relations.* Four words are presented; the examinee must tell how the first three are similar to each other but different from the fourth word.
>
> Quantitative Reasoning (Crystallized Abilities)
>
> *Quantitative.* Manipulatives, pictures, and verbal descriptions provide problems that require use of numerical reasoning, concepts, and computation.
>
> *Number Series.* Number Series requires examinees to discern a pattern in a series of numbers and to use that pattern to predict the next two numbers in the series.
>
> *Equation Building.* Pictorially presented groups of numbers and operation signs (+,x,=, etc.) are arranged into balanced equations.
>
> Abstract/Visual Reasoning (Fluid-Analytic Abilities)
>
> *Pattern Analysis.* Form board or block-design-type tasks are solved, depending on the examinee's age and ability levels.
>
> *Copying.* Geometric designs are reproduced with blocks or paper and pencil depending on examinee's age and ability levels.
>
> *Matrices.* Visually presented analogies, similar to Raven's Progressive Matrices, are solved.
>
> *Paper Folding and Cutting.* Pictures are selected as representations of how folded and cut pieces of paper would look unfolding.
>
> Short-Term Memory
>
> *Bead Memory.* Pictorially presented bead patterns are reproduced by placing beads on a stick.
>
> *Memory for Sentences.* Orally presented sentences are repeated from memory.
>
> *Memory for Digits.* Memory for Digits is a simple digits forward and digits reversed task.
>
> *Memory for Objects.* Pictures objects are recalled in the sequence of presentation.
>
> Keith et al. (1988, p. 255-256)

Within broad limits the examiner is free to choose from a variety of test combinations. Subject age, ability and examiner preferences/objectives are primary considerations. Figures 2, (p. 6) and 6, (p. 14) in the Guide For Administering And Scoring The Fourth Edition (Thorndike, Hagen, & Sattler, 1986a) outline age and ability ranges of each test. It is evident that with any particular subject the examiner is free to use up to thirteen of the fifteen available tests. A complete battery consists of eight to thirteen tests depending upon the chronological age and ability level of the subject. Several special purpose abbreviated batteries are recommended and these are displayed in figure 11 (p. 36) of the *Administration Guide*. These batteries range from four to eight tests in length.

Various scores, therefore, are available with the SBIV. Raw and Standard Age Scores for each of the 15 test are obtainable. Four cognitive area scores are calculable. A composite of all four or any combination of the four area scores can be obtained. A profile on all 15 tests (based on Standard Age Scores) is possible.

Each of the fifteen tests has a mean of 50 and standard deviation of 8. Area and Composite scores have a mean of 100 and standard deviations of 16. The latter, is consistent with earlier editions.

Test Administrators

All examiners were students enrolled in a graduate intellectual assessment course. Test procedures were taught according to the guidelines outlined by Thorndike, Hagen and Sattler (1986a). All test protocols were checked for administration and scoring accuracy by University of Alberta Faculty or their trained assistants.

The Facility

The Education Clinic is a well established service operated under the auspices of the Department of Educational Psychology, Faculty of Education, University of Alberta. Its primary function is to train graduate students in School Psychology and Counseling programs. Referrals are accepted from schools, parents, individuals and various agencies. Services are free of charge.

Participants

The present sample consists of three hundred and seventy one individuals who were between the ages of 3 (no two year olds were among the population tested) and 23 years. This group was a subsample of all individuals who were administered the SBIV at the University of Alberta, Education Clinic between January, 1987 and April, 1988. A total of four hundred and sixteen individuals were administered the SBIV during this period. Twenty-one of these individuals were older than 23 years of age and hence fell outside the SBIV standardization sample age range. Twenty-four protocols failed to meet the full basal criteria outlined by Thorndike, Hagen and Sattler (1986a) and were excluded for this reason.

The four hundred and sixteen individuals initially tested with the SBIV represented approximately one-third of all individuals who were administered a standardized intellectual assessment during the above noted period. Two other intelligence tests are routinely administered at the Education Clinic. Assignment to any group is arbitrary and there were no strict guidelines used, other than those pertaining to the age requirements of each instrument.

The nature of the sample is outlined in tables 3.0 to 3.8 and analyzed using the Statistical Package for the Social Sciences (SPSS, 1988). This information can be compared to the standardization data that is provided in the *Technical Manual* (Thorndike, Hagen & Sattler, 1986b, p. 22-26). The average age of the clinic sample is 8.9 years. When the clinic sample is divided into age categories a reasonable distribution is obtained. So too is the representation of gender groups. Caucasians and urban dwellers are somewhat more represented in the clinic sample. The highest level of education and parental occupation is selected when two parents are involved. However, the overrepresentation of higher socioeconomic status parents and participants who possess greater than high school education is less than the standardization sample. Referral reasons are diverse. The majority of referral reasons fall into two general categories, parental interest and learning concerns. Importantly both mentally handicapped and gifted populations are represented.

In summary, there are several reasons supporting the use of this clinic sample. First, the quality of test administration procedures and scoring is overseen by highly qualified experts. Secondly, the Education Clinic provides service to a wide ranging clientele that is, in general, similar to both the general population and individuals routinely seen by educational psychologists. An important task is, however, to determine if the factorial structure for a clinical population is the same as the standardization sample. Thirdly, the logistical effort required to obtain completed protocols is onerous. Finally, the Education Clinic offered a sample that is easily accessible.

TABLE 3.0

Age Distribution for the Entire Sample

Age	Frequency	Percent
3-0 to 3-11	9	2.4
4-0 to 4-11	28	7.5
5-0 to 5-11	49	13.2
6-0 to 6-11	44	11.9
7-0 to 7-11	42	11.3
8-0 to 8-11	25	6.7
9-0 to 9-11	28	7.5
10-0 to 10-11	32	8.6
11-0 to 11-11	20	5.4
12-0 to 12-11	23	6.2
13-0 to 13-11	15	4.0
14-0 to 14-11	18	4.9
15-0 to 15-11	8	2.2
16-0 to 16-11	2[illegible]	5.7
17-0 to 17-11	4	1.1
18-0 to 18-11	0	0
19-0 to 19-11	0	0
20-0 to 20-11	1	.3
21-0 to 21-11	1	.3
22-0 to 22-11	1	.3
23-0 to 23-11	2	.5
Total	371	100.0
Mean	8.9	SD 4.0

TABLE 3.1

Age Distributions by Age Groups

Age Group	Frequency	Percent	Mean	Standard Deviation
(1) 3 - 6 yrs	130	35.0	5.0	0.9
(2) 7 - 11 yrs	147	39.6	8.8	1.4
(3) 12 - 23 yrs	94	25.3	14.4	2.4
(4) 3 - 23 yrs	371	100.0	8.9	4.0

TABLE 3.2

Sex Distributions By Age Groups

Age Group	Male		Female		Total	
	#	%	#	%	#	%
3 - 6 yrs	75	57.7	55	42.3	130	100.00
7 -11 yrs	84	57.1	63	42.9	147	100.00
12 -23 yrs	54	57.4	40	42.6	94	100.00
3 -23 yrs	213	57.4	158	42.6	371	100.00

TABLE 3.3

Ethnic Distributions By Age Groups

Age Group	Caucasian		Native		Asian		Hispanic		Black		Other	
	#	%	#	%	#	%	#	%	#	%	#	%
3 - 6yrs	118	90.8	2	1.5	5	3.8	1	.8	1	.8	3	2.3
7 - 11yrs	134	91.2	3	2.0	2	1.4	3	2.0	1	.7	4	2.7
12 - 23yrs	71	75.5	5	5.3	4	4.3	1	1.1	1	1.1	12	12.8
3 - 23yrs	323	87.1	10	2.7	11	3.0	5	1.3	3	.8	19	5.1

TABLE 3.4

Parental Occupation Distributions By Age Groups

Occupation	3-6 yrs #	3-6 yrs %	7-11 yrs #	7-11 yrs %	12-23 yrs #	12-23 yrs %	3-23 yrs #	3-23 yrs %
Managerial/Professional	66	50	51	34.7	27	28.7	144	38.8
Technical Sales	20	15.4	19	12.9	13	13.8	52	14.0
Service Occupations	10	7.7	31	21.1	22	23.4	63	17.0
Farming/Forestry	3	2.3	6	4.1	5	5.3	14	3.8
Precision Production	7	5.4	11	7.5	6	6.4	24	6.5
Operators/Fabricators	10	7.7	12	8.2	11	11.7	33	8.9
Unemployed	0	0	6	4.1	3	3.2	9	2.4
Unknown	9	6.9	9	6.1	6	6.4	24	6.5
Other	5	3.8	2	1.4	1	1.1	8	2.2
TOTAL	130	35.0	147	39.6	94	25.3	371	100.0

TABLE 3.5

Parental Education Distributions By Age Groups

Age Groups	College Grad or > #	College Grad or > %	1-3 yrs College #	1-3 yrs College %	HS Grad #	HS Grad %	< HS #	< HS %	Unknown #	Unknown %
3 - 6 yrs	61	46.9	10	7.7	23	17.7	5	3.8	31	23.8
7 - 11 yrs	47	32.0	11	7.5	25	17.0	17	11.6	47	32.0
12 - 23 yrs	22	23.4	4	4.3	18	19.1	14	14.9	36	38.3
3 - 23 yrs	130	35.0	25	6.7	66	17.8	36	9.7	114	30.7

TABLE 3.6

Parental Marital Status Distributions By Age Groups

Age	Married #	Married %	Divorced #	Divorced %	Separated #	Separated %	Single #	Single %	One Parent Deceased #	One Parent Deceased %	Unknown #	Unknown %
3 - 6	122	93.8	4	3.1	3	2.3	1	.8	0	0	0	0
7 - 11	130	88.4	7	4.8	4	2.7	1	.7	4	2.7	1	.7
12 - 23	83	88.3	4	4.3	2	2.1	0	0	3	3.2	2	2.1
3 - 23	335	90.3	14	4.0	9	2.4	2	.5	7	1.9	3	.8

TABLE 3.7

Geographical Distribution By Age Groups

Age Groups	299,999 or >		100,000 to 299,999		25,000 to 99,999		2500 to 24,999		<2300	
	#	%	#	%	#	%	#	%	#	%
2 - 6 yrs	91	70	2	1.5	9	6.9	19	14.6	9	6.9
7 -11 yrs	87	59.2	1	.7	15	10.2	25	17.0	19	12.9
12 -23 yrs	64	68.1	0	0	6	6.4	11	11.7	13	13.8
2 -23 yrs	242	65.2	3	.8	30	8.[illegible]	55	14.8	41	11.8

TABLE 3.8

Reason for Referral Distribution By Age Groups

Age	Self/ Parent Interest		Learning Concerns		Emotional/ Develop Concerns		Giftedness		Mentally Deficient		Special class/ School Placement	
	#	%	#	%	#	%	#	%	#	%	#	%
2 - 6 yrs	92	70.8	16	12.3	4	3.1	8	6.2	0	0	10	7.7
7 -11 yrs	48	32.7	61	41.5	10	6.8	8	5.4	4	2.7	16	10.9
12 -23 yrs	17	18.0	30	32.0	2	2.2	11	11.7	11	11.7	23	24.5
2 -23 yrs	157	42.3	107	28.8	16	4.3	27	7.3	15	4.0	49	1[illegible].2

Cases Deleted from the Sample

A review of the deleted cases is also necessary. The purpose of this review is to determine if they differ substantially from the retained sample. If significant or persistent differences are present it would be important to investigate possible explanations or trends in these protocols.

The average age of the deleted cases (excluding those over 23 years) is 8.5 years. This is highly similar to the retained samples mean of 8.9 years. Such a difference is small and indicates that no major age deviations exist between retained and deleted cases.

The sex distributions of both groups revealed that they are split proportionally in a similar fashion. However, the male/female composition is reversed. Females are slightly under-represented in the retained sample (i.e., 42.6 percent) and somewhat over-represented in the deleted sample (i.e., 55.6 percent). Similarly males are slightly over-represented in the retained sample (i.e., 57.4 percent) and somewhat under-represented in the deleted sample (i.e., 44.4 percent). These differences are slight and believed to be inconsequential.

The ethnic breakdown for both samples is essentially the same. Approximately 89 percent of the deleted cases are caucasion. This is similar to the 87 percent observed for the retained sample.

Differences in the parental occupational backgrounds are slight. Where differences are present the small number of deleted protocols is likely the influencing factor. This is also true for the small differences observed when parental education levels between the two groups are contrasted.

Parental marital status differences are also small. The vast majority of parents in both groups are married. For the retained sample 90.3 percent are married. The figure for the deleted sample is 86.7 percent. Other differences are small and are likely influenced by the small number of deleted protocols.

Differences in community size of origin for retained and deleted cases are small. The majority of cases in both groups are living in large urban centres. Differences between categories are seldom greater than five percentage points.

The referral reasons are somewhat different for the two groups. A large proportion of individuals in the deleted sample (i.e., 68.9 percent) are of the parental interest variety. This number is significantly different than that obtained for the retained sample (i.e., 42.3 percent). Learning concerns are more prominent in the retained sample (i.e., 28.8 percent vs. 6.7 percent) Therefore, relative to the deleted sample, the retained sample is more diverse and appears to be a more accurate reflection of the general population. In summation, a comparison of the retained and deleted samples is largely favorable and suggests that the retained and deleted samples are not systematically biased in any areas that are of concern.

Age Categorizations

As noted earlier, the Binet scales are characterized by their adaptive testing and age scale formats. Consequently, individuals at various ability and age levels are evaluated on material of varying degrees of similarity. Compounding the complexity of this issue is that several tasks are thought to involve different abilities at various developmental levels (Keith et al., 1987; Vernon, 1987). Age groupings are typically employed prior to computational work to account for this methodological concern. Such an approach is adopted here and each age group is treated as if it is a separate sample.

Fritzke (1988) criticized the age groupings that are used by Thorndike, Hagen and Sattler (1986b). The former's scheme is much more elaborate and complicated. Fritzke's logic is sound. However, the purpose of this study is to examine the theoretical and empirical data as set forth by Thorndike, Hagen and Sattler (1986b). As noted earlier no 2 year olds are present in the sample simply because of the sample that is utilized. Therefore, the age groupings selected by the test authors are, for the most part, employed.

Dealing With Missing Data

The SBIV test format is statistically troublesome for a second reason. It is possible for both Area and by default Composite scores to be comprised of various numbers and types of tests. This is problematic for both comparative and analytic purposes.

Timm (1970) and Gleason and Staelin (1975) have examined the problem of missing observations in multivariate data. Both studies have reviewed various, commonly accepted methods and techniques for dealing with this problem. Both studies are supportive of multiple regression techniques, although Gleason and Staelin (1975) propose an alternative method that is equivalent to multiple regression under certain more restrictive conditions. Therefore, multiple regression techniques are the preferred way of dealing with missing data.

There are, however, limitations of such prediction techniques. A minimum number of values are required in each of the to be predicated tests. Otherwise the risk of error is likely to surpass the potential gains obtained by having an increased number of entries for each test. A decision was made that each retained test should not have more than 35 percent of their entries missing. The number and percentage of missing entries for each subtest by age groupings are outlined in Table 3.9.

Using the 35 percent plus decision seven tests are excluded from the 3 to 6 year old analysis. These are Verbal Relations, Number Series, Equation Building, Matrices, Paper Folding and Cutting, Memory for Digits, and Memory for Objects. Three of these tests have age ranges outside this category. Verbal Relations, Equation Building, Paper Folding and Cutting are only available to subjects eight year of age and older. Number Series, Matrices, Memory for Digits, and Memory for Objects are only available for subjects five years and over. These latter tests also have a minimum entry level of I that further restricts their availability for use.

TABLE 3.9

Number and Percentage of Missing Cases by Subtest for Each Subsample and Entire Sample

Age	3 to 6 yrs #	3 to 6 yrs %	7 to 11 yrs #	7 to 11 yrs %	12 to 23 yrs #	12 to 23 yrs %	3 to 23 yrs #	3 to 23 yrs %
Vocabulary	0	0	0	0	0	0	0	0
Comprehension	0	0	0	0	0	0	0	0
Absurdities	0	0	3	2	21	22.3	24	6.5
Verbal Relations	130	100	146	99	55	58.5	331	89.2
Pattern Analysis	0	0	0	0	0	0	0	0
Copying	3	2.3	18	12.2	40	42.6	61	16.4
Matrics	118	90.8	60	40.8	14	14.9	192	51.8
Paper Folding and Cutting	130	100	145	98.6	51	54.3	326	87.9
Quantitative	0	0	0	0	0	0	0	0
Number Series	125	96.2	67	45.6	23	24.5	215	58
Equation Building	130	110	143	98.6	67	71.3	342	92.2
Bead Memory	0	0	0	0	0	0	0	0
Memory for Sentences	0	0	0	0	0	0	0	0
Memory for Digits	84	64.6	23	15.6	6	6.4	113	30.5
Memory for Objects	92	70.8	30	20.4	8	8.5	130	35.0

Five tests are excluded from the 7 to 11 year old analysis. These are Verbal Relations, Matrices, Paper Folding and Cutting, Number Series and Equation Building. Verbal Relations, Paper Folding and Cutting, and Equation Building are available only to individuals eight years of age or older. In addition, the latter three tests have a minimum entry level of M. Number Series and Matrices, while available to the 7 to 11 year old subsample have a minimum entry level of I.

Four tests are excluded from the 12 to 23 year old analysis. These are Verbal Relations, Copying, Paper Folding and Cutting, and Equation Building. Copying (along with absurdities) has a upper chronological age range of nine years. Verbal Relations, Paper Folding and Cutting, and Equation Building all have minimum entry levels of M.

A second issue related to missing data is the possibility that the SBIV increases the opportunity for examiners to overlook compulsory subtests. This is because the SBIV, relative to the Wechsler Scales, offers an increased choice in core and optional tests. Examiners are required to exercise greater care in ensuring required tests are selected. The SBIV authors are able to rightfully counterbalance this claim by stating that the compulsory, substitutional and optional tests are clearly specified in the *Administration Manual*. In addition, the data presented earlier suggests that this is likely not a major problem when test administrators are adequately trained.

In summary, the propensity for missing optional test data is likely a serious liability of the SBIV only when it is used for comparative or research purposes. The possibility of different analyses using different combinations of tests is moderately high. Steps should be routinely taken to ensure that all available tests are completed and that populations are equivalent in both age and ability prior to making statements of equivalencies.

The problem of missing test data is more cumbersome when the factor structure of the SBIV is being considered. This is because this analysis relies on the intercorrelations of tests. Extreme correlation values may arise particularly when specific pairs of tests are given to a select part of the sample. Thorndike, Hagen and Sattler (1986b) use median rather than mean correlations "to minimize the weight given to extreme values of ... correlation" (p. 52). It can, however, be argued that the pooling of median correlations clouds the interpretative process because it allows for negative eigenvalues. More appropriate and acceptable practices include pooling weighted covariances to compute average correlations and using Fisher's Z_r. Suffice it to say that having subjects complete as many subtests as possible minimizes but does not resolve the problem of missing data on the SBIV.

A second consideration is the type of tests that are selected to predict missing values. The six core tests are utilized. three of these have a verbal aspect to them. However, the fact that these six subtests are selected off sets this bias to some degree. Therefore, the utilization of the core tests is viewed as the best available option.

In summation, multiple regression procedures are viewed as the technique of choice for predicting missing data entries. Estimates of missing data entries are obtained in the following fashion. Regression equations are calculated separately for each age grouping utilizing the six compulsory tests. Each missing test score is then calculated in a step-wise fashion. Only those core tests making a statistically significant (.05) contribution to the prediction of test scores are retained in the multiple regression equation. Table 3.10 outlines the regression equations at the .05 level of significance for each predicted test score by age categories.

Analytic Techniques

Five major factor analytic procedures are employed in this study. The purpose of employing these methods is twofold. First, it is to determine if a consistent factor pattern emerges. If a similar or supportive pattern emerges from these diverse procedures the results can be considered more robust. Secondly, it is to determine if one method provides a more psychologically meaningful explanation of the data. These five procedures are; unrotated principal components, rotated principal components, Procrustes, multiple-group and hierarchical factor analysis.

TABLE 3.10

Regression Equations for Missing Values by Subsample

3 to 6 yrs	COPY = 24.6387 + .31078 x PATTERN ANALYSIS SCORE + .20004 x MEMORY FOR SENTENCES SCORE, r^2 = .228
7 to 11 yrs	ABSURDITIES 12.32108 + .44212 x VOCABULARY SCORE + .32978 x PATTERN ANALYSIS SCORE, r^2 = .491
	COPY = 14.80483 + .35214 x PATTERN ANALYSIS SCORE + .27898 x QUANTITATIVE SCORE, r^2 = .278
	MEMORY FOR DIGITS = 15.51690 + .44193 x MEMORY FOR SENTENCES SCORE + .26758 x QUANTITATIVE SCORE, r^2 = .542
	MEMORY FOR OBJECTS = 19.02326 + .35823 x PATTERN ANALYSIS SCORE + .27545 x VOCABULARY SCORE, r^2 = .409
12 to 23 yrs	ABSURDITIES = 8.63629 + .38980 x COMPREHENSION SCORE + .40986 x PATTERN ANALYSIS SCORE, r^2 = .699
	MATRICES = .43548 + .26433 x QUANTITATIVE SCORE + .20018 x BEAD MEMORY SCORE + .31641 x PATTERN ANALYSIS SCORE + .21856 x VOCABULARY SCORE, r^2 = .711
	NUMBERS SERIES = -.53886 + .33166 x QUANTITATIVE SCORE + .27897 x PATTERN ANALYSIS SCORE + .27455 x COMPREHENSION SCORE + .14399 x BEAD MEMORY, r^2 = .834
	MEMORY FOR DIGITS = 13.07447 + .31610 x MEMORY SENTENCES SCORE + .22146 x BEAD MEMORY + .21797 x QUANTITATIVE SCORE, r^2 = .589
	MEMORY FOR OBJECTS = 21.27530 + .29126 x VOCABULARY SCORE + .26889 x MEMORY FOR SENTENCES, r^2 = .377

Unrotated principal components. Transforming raw data matrices to either correlation or covariance matrices is customary prior to conducting a principal components analysis. The most common approach is to use correlation matrices based upon ordinary Pearson product-moment coefficients. Correlation coefficients are used to address the fact that frequently one is dealing with variables that have different units and scales (Dillon & Goldstein, 1985; for detailed discussions on the advantages and disadvantages of various correlational techniques as well as differences see Gorsuch, 1983; Nunnally, 1970; Rummel, 1970). "In computing a correlation coefficient between two variables, differences due to both the mean and the dispersion of the variables are removed. Thus, the transformation makes the variables directly comparable" (Dillon & Goldstein, 1985, p. 26).

Prior to discussing this and other procedures a brief diversion is necessary to highlight the impact the nature of the sample can have on the obtained findings. Age and ability ranges are primary considerations. Comrey (1978) stresses that when ability tests are given to children who range widely in age a general factor may emerge simply because most older children do better than younger children. As a results the "correlations between all tests will be very high, and a single factor will account for most of the variance. An unwary investigator might call this a general factor, but in reality it is only a maturation factor" (p. 650). This problem is lessened when age-related samples are employed.

A second consideration that is highlighted by Comrey (1978) is that populations that are very selective reduce the variance on general factors, thereby highlighting group factors. Gorsuch (1983) also addresses this complex issue:

> Because selection usually reduces correlations, it appears that selection is more likely to eliminate factors than to produce them. In addition, the reduction in the correlations reduces the length of the variable vectors and thus makes them more subject to error fluctuations. For reasons such as these, *the usual recommendation is to randomly sample as broadly as possible so that individuals are included from all of the population to which the results might be generalized* (p. 346)(italics in original).

As noted earlier, the inclusion of subjects of all ability levels in each subsample is believed to offset, to some degree, the special nature of the Clinic population. Nevertheless, the fact that it is derived from a Clinic population should not be ignored. An additional concern identified by Comrey (1978) is the problem of outliers or unusual scores that can significantly influence correlation matrices. Outliers are potential sources of spurious common factor variance that can influence data interpretation. As discussed earlier, outliers are not believed to exert a significant influence in the obtained data. Comrey, however, (1978) cautions researchers to inspect all scores that are out of the ordinary for possible explanations and exclusion if necessary (Carroll, 1985). Several other issues of direct interest to prospective factor analytic investigators are raised in this article. In addition, he reviews a number of textbooks that are often referenced in factor analytic works. Several of these topics are discussed here as they become relevant to data interpretations.

Carroll (1985) has summarized the two major goals of factor analysis in the study of intelligence and intellectual abilities. There goals are: "(a) to identify a minimum number of latent variables or "factors" in a set of data that will most parsimoniously, and from the standpoint of psychological interpretation most intelligibly, account for covariation among the variables in the data, and (b) to determine the "structure" of these factors, that is to reveal any hierarchical arrangement of the factors such that some factors are more general than others" (p. 25). Accordingly, Carroll (1985) emphasizes that factors identified in one study as being basic may , in others, be viewed as a composite of more primary abilities.

Rummel (1970) highlights the fact that factor analysis is "a general scientific method for analyzing data" (p. 13). He notes that it is possible to factor almost any type of matrix. A second aspect of his statement is that there are many different factor analytic techniques. Dillon and Goldstein (1985), as such, have referred to factor analysis as a "family" of techniques. Under certain conditions, however, the results of various techniques are likely to produce similar results (See Go...ch 1983 for a detailed discussion on replication and invariance of factors).

A basic distinction between factor analytic methodologies is the assumption made about common, unique and total variance. The primary purpose of principal components analysis is to specify the fewest number of factors that account for the largest proportion of total variance. No attempt is made to separate common and unique variance. Principal components are extracted in order of decreasing magnitude. The first component is the linear combination of variables that account for the largest proportion of total variance. The second principal component is uncorrelated with the first and accounts for largest amount of total remaining variance. It is possible to continue this process to the point where there are as many factors as there are variables. However, the procedure is usually stopped when a large proportion of the data is accounted for and successive factors would add little to the total variance and/or understanding of the data (Dillon & Goldstein, 1985; Gorsuch, 1983).

Unlike principal components analysis the common factor analytic model assumes that a variable is divided into common and unique parts." The *'common part'* of a variable is that part of the variable's variation that is shared with the other variables, whereas the *'unique part'* of a variable is that part of the variable's variation that is specific to that variable alone" (Dillon & Goldstein, 1985, p. 56). The specification of what common factors are present in a variable is the chief interest of the common factor model and this has additional implications for variable specification. Carroll (1985) writes:

> Related to this determination of the *'communality'* of a variable, that is, the proportion of its variance that is accounted for by the common factors identified in the data set. The complement of the communality is the (squared) *'uniqueness'* of the variable. If the reliability of the variable is known or can be estimated (by commonly accepted techniques outside of factor analysis), the uniqueness can be divided into the variables *specificity* and its *error variance*. These quantities, expressed as proportions, constitute information that is useful in further studies of the variables" (Italics in original p. 25).

It is evident from the above that principal components analysis is variance oriented. The common factor model is, however, primarily concerned "with the amount of each variable's variance that is shared with other variables, and therefore the common factor model is covariance oriented" (Dillon & Goldstein, 1985, p. 71). As a result the observed scores need to be adjusted to reflect only their covariant parts. These adjusted values (communalities) are then used to replace the unities in the diagonals of the "to-be-factored" correlation matrix. The common factor matrix is estimated by subtracting an estimate of r from the observed correlation matrix. The communality problem arises because of the need to estimate appropriate diagonal values.

Several methods are available to estimate communality values (Dillon & Goldstein, 1985). One method used is to select the largest correlation regardless of sign each variable has with any other variable. Dillon and Goldstein (1984) note that this procedure is more appropriate for problems involving a large number of correlations. Other methods are available using only a few of the correlations in the correlation matrix (i.e., placing unities in diagonal such as in the principal components case). However, certain methods can lead to Heywood cases (i.e., situations where communality estimates are greater than one). When this occurs the communality is usually set to 0.99 or 1.0. A communality estimate of 1.0 indicates that the variance of a variable is entirely explained by available common factors. Conversely, a communality of 0.0 is present when the variable does not correlate with any other variable in the matrix.

Other methods that use the entire correlation matrix are available. A commonly employed and frequently recommended method is to use the squared multiple correlations (SMC). "Thus for variable 1 in a group of 20 variables, this would be the SMC for variables 2 through 20 with variable 1 (Nunnally, 1967, p. 353). This method was proven by Guttman (1956) to be the lowest boundary for communality. Nunnally (1967) has summarized the difficulties and advantages of this approach:

> There are two major difficulties with the use of SMCs as estimates of communalities. First, they obviously determine one type of common variance, the variance that a particular variable has in common with the other variables in a matrix...(as opposed to a set of common hypothetical factors)... A second problem with the use of SMCs is that they do not reproduce the expected results in artificial problems where stated sets of factor loadings are used to generate tables of correlations... The SMCs..., (however), have the advantages of being (1) unique, (2) directly obtainable on computers, and (3) definitive of at least one type of common variance (p. 354).

A second approach that uses the entire correlation matrix dates back to Thurstone. This is the iterative approach. This procedure begins by estimating the communalities and placing this estimate in the diagonals. This matrix is then factored. The "sums-of-squares of factor coefficients for a *predetermined number of factors* are used as a new set of communalities" (Dillon & Goldstein, 1985, p. 73 italics in original). This new correlation matrix is then factored and the procedure is repeated until changes in the correlation coefficients are small. The final loadings are accepted as the best estimate of communalities.

Nunnally (1967) has identified the two major difficulties with iterative approaches. He notes that different initial communality estimates can lead to different communality solutions and factor loadings. Secondly, iterations can lead to Heywood cases for some variables. Other approaches are available but Nunnally (1967) stresses that in the final analysis no method is problem free. He writes that fortunately research indicates that communality estimates closely approximate analyses that use unities in the diagonals. "Thus even when the use of unities in the correlation matrix tends to confound common variance and unique variance, the confounding usually is slight" (Nunnally, 1967, p. 353).

In summation, principal components analysis is concerned with basic dimensions of the data. Assumptions are not made about common aspects of these dimensions. Common, specific and random error variances are mixed (Rummel, 1970).

The choice of which method to use is considered one of the first important decisions to make in factor analytic research (Carroll, 1985). Principal components is often advocated because it (a) derives a small set of linear combinations of variables that account for most of the data (Dillon & Goldstein, 1985) (b) is conceptually and computationally simplier (Velicer, Peacock & Jackson, 1982) (c) produces essentially the same results as the common factor model (Velicer, Peacock & Jackson, 1982) (d) avoids the basic indeterminacy problem. Dillon and Goldstein (1984) write:

> We will shortly see that there is a basic indeterminacy associated with the common-factor model: In order to determine the number of common factors precisely, we need to know the variance of each variable common to the other p - 1 variables; on the other hand, until the number of common factors is known, that part of a variable's variance that is shared with the other variables cannot be determined (p. 56).

The common factor model is, however, not without its advocates. Dillon and Goldstein (1985) stress that "common factor-analytic techniques can better serve the functions of searching the data for qualitative and quantitative distinctions and, especially, testing *a priori* hypotheses about the number of common factors underlying a set of data or the character of common factors" (Dillon & Goldstein, 1985, p. 56). Carroll (1985) also advocates the use of the common factor model because it allows for the distinction between common and unique variance. He believes that this advantage allows for a more accurate understanding of the "true factor structure of a set of data" (p. 33).

Velicer and Jackson (1990) have reviewed the claims of superiority which have been advocated by both component and common factor proponents. The focus on explicit mathematical models and available empirical evidence. Velicer and Jackson provide several convincing reasons for equivalencies as well as the superiority of the component model under various circumstances (1990, p. 110).

As noted above a pivotal point in the choice of techniques debate is the issue of factor-score indeterminacy. McDonald and Mulaik (1979) provide a relatively nontechnical review of this issue. They offer the following definition of this problem:

> Factor-score indeterminacy refers to the fact that the common and unique factor scores in the common factor model are not uniquely determined by the observed variables whose correlations they explain, since in general the multiple correlation between a common or unique factor and the observed variables is less than unity (p. 297).

These authors highlight that this problem centres around the fact that factor scores do not have a unique mathematical definition. Consequently, infinite sets of numbers can be generated that will match a given set of observations. Therefore, this problem "concerns the inability of a finite set of observed variables in an exploratory factor analysis to determine unambiguously what attribute of the individuals the factor variable represents" (McDonald & Mulaik, 1979, p. 298). They further note that many researchers are not concerned about factor score indeterminacy. This is because these researchers see little relationship between this issue and their primary research goal of looking at the common aspects of variables with high loadings. However, as noted above the infinite number of mathematical constructions of factor variables is related to the number of possible interpretations of a common factor.

McDonald and Mulaik (1979) emphasize that the problem of factor indeterminacy is relatively more pronounced when factor analysis is used in an exploratory fashion. When variables are selected on preconceived notions of certain attributes there is less uncertainty regarding what variables make up the common components. However, because few variables have an agreed upon domain it is still possible, using different domains or portions of the same domain to come up with distinct common attributes.

McDonald and Muliak (1979) conclude, therefore, that this problem of common factor models is not sufficient for justifying different methods of analysis such as principal components. They emphasize that one cannot guarantee a component score "of a core set of variables uniquely determines a corresponding score in an infinite domain from which the core set has been drawn" (p. 305). Therefore, McDonald and Mulaik (1979) believe that this problem must be restated in "terms of behavior domains and in terms of the relation of core components to domain components" (p. 305). Their interpretation of the determination of common factors is summarized below:

> (I)f the core set...(of variables)...can be given the same factor loadings on a factor when analyzed alone or in the context of an infinite domain of variables, then there is just one factor variable in the domain that is a possible factor variable of the core

> set. If consistent factor loadings cannot be found, there is no factor variable in the domain that is a possible factor of the core set. In the latter case, alternative subdomains of variables may contain alternative possible factors of the core set (McDonald & Mulaik, 1979, p. 297).

A second issue of indeterminacy is that even though an unrotated factor solution may be specific to a given data matrix, the number of mathematically equivalent rotated solutions is infinite (Thurstone, 1947; Hill, Reddon & Jackson, 1985). Gorsuch (1983) recommends that an appropriate rotation is possible if certain criteria are examined. These criteria are substantive, orthogonality and mode of rotation.

Thurstone's (1947) simple structure is probably the most well known and accepted substantive criteria. "Simple structure involves rotating the factors such that each better defines a separate cluster of highly interrelated variables and is as specific to this cluster as possible" (Gorsuch, 1983, p. 170). Other substantive criteria are available.

A second criteria identified by Gorsuch (1983) is orthogonality. Here rotated factors are uncorrelated. Kaiser's (1958) varimax is an orthogonal procedure and its results parallel those of simple structure (Hill, Reddon & Jackson, 1985). However, Kaiser (1958) argues that the property of factor invariance inherent in the normal varimax solution is of greater significance than the property of simple structure.

Gorsuch's final criterion is whether or not the rotation should be graphically or analytically based. Positioning the axis according to visual simplicity is the basis of the graphic approach. Graphical rotation is labor intensive and impractical when the number of variables is large. Analytic techniques are designed for analyses involving large numbers of complex variables. Analytic rotations position the initial factors according to a mathematical criterion. "Orthogonal rotation may or may not be done prior to oblique rotation. Often, however, it is desirable to have an orthogonally rotated solution for comparing oblique results" (Gorsuch, 1983, p. 171). The different solutions can then be compared according to both theoretical and clinical usefulness.

Oblique rotations allow for the correlation between clusters that are interrelated. If no correlation exists between variables oblique procedures will lead to orthogonal factors.

If, however, the variables are related, however, oblique procedures provide additional information on the nature of this relationship. Oblique procedures are required to determine if intelligence is hierarchically organized. The three orthogonal rotations employed in this study are outlined below.

Exploratory Factor Analysis

Rotated Principal Components. Analytic rotations, as noted earlier, are designed to facilitate the positioning and interpretation of factors. Gorsuch (1983) has identified several reasons why analytic techniques are favored over visual rotational methods. "First, the replicability and quality of analytic simple structure rotations can be easily investigated... Second, visual rotation can only be considered objective if it is carried out without any knowledge of the identity of the variables" (p. 189). Gorsuch (1983) warns that rotating without prior knowledge of the nature of variables can lead to the acceptance of chance relationships. Finally, Gorsuch (1983) emphasizes logistic considerations:

> In a large factor analysis, the time necessary to rotate the factors to visual simple structure is extensive. It will often take the investigator several hours each day for five or six weeks to obtain a rotated position which has a high likelihood of being the best simple structure, not to mention the time necessary for computations. For the same-sized matrix,...(several procedures outlined below)...will take from less than a minute to, at most, an hour or so of computer time (p. 189).

As stated, there are two broad ways in which factor axes can be rotated. Orthogonal rotations are perpendicular or uncorrelated after rotation. Oblique rotations are less restrictive and are only perpendicular after rotation if that is the true nature of the data. Oblique rotations allow for independent rotation of each factor axes.

The distinctive properties of orthogonal rotational procedures have been summarized by Dillon and Goldstein (1985):

> 1. Factors resulting from the orthogonal rotation of principal components will remain statistically uncorrelated; that is, the cosine of the angle between rotated factors is zero.
> 2. Any orthogonal rotation method will not alter the values of the communality estimates... However, the properties of a variable's variance accounted for by a given factor will be different.
> 3. Though the total amount of variance accounted for by the common factors does not change with orthogonal rotation, the percentage of variance accounted for

> by an individual factor will, in general, be different. Thus, in the orthogonally rotated factor pattern matrix, no significance is attached to factor order. Obviously, the percentage of common variance accounted for by a common factor will also change after orthogonal rotation (p. 92).

Three commonly used orthogonal rotational procedures are employed in this study. These are quartimax, varimax and equimax. They each differ on how simple structure is defined.

The quartimax method is the earliest attempt to define Thurstone's simple structure by more precise mathematical criteria. "It is based on one of the salient requirements of simple structure - that the variance of loadings in each row of the factor matrix should be as large as possible" (Nunnally, 1967, p. 332). When this occurs, the amount of variance a variable has on other factors is, by default, small. Therefore, each variable is loading mainly on a single factor Gorsuch (1983) states that computationally this is done by either minimizing the cross-products between squared loadings on a given variable or maximization of the fourth powers of the factor loadings (also see Harman, 1967, pp. 298-304).

A well recognized shortcoming of the quartimax method is its tendency to produce a General factor (Gorsuch, 1983, 1983; Harman, 1967; Nunnally, 1967). As Nunnally (1967) writes..."(o)ne can see why this would be so, because the criterion would be perfectly satisfied if all variables had a loading on only one factors" (p. 232). Gorsuch (1983) is also critical of this approach. He stresses that ...."(w)hile each variable's complexity is kept near the minimum, the complexity of the first factor is maximized rather than minimized (p. 1). Therefore, while the quartimax method is often used these deficiencies, in terms of simple structure, must be taken into account.

Kaiser (1958) defined an analytic criterion for rotation that is generally considered to be a modification of the quartimax method that more closely approximates simple structure. The varimax method, however, is designed to reduce the complexity of columns (factors) rather than rows (variables). Thus, varimax procedures are designed to maximize the variance across factors (Gorsuch, 1983, 1983; Harman, 1967). Nunnally (1967) writes:

> Rather than maximize the sum of variances of rows in the matrix,...(Kaiser's)...*varimax* method maximizes the sum of squared loadings in the columns of the factor matrix. In each column of the matrix, this tends to produce some high loadings and some loadings near zero, which is one aspect of simple structure (p. 332).

Therefore, when a varimax solution is employed, the opportunity for a general factor is minimized. Because of this Gorsuch (1983) writes "...*(v)arimax* cannot be used if the theoretical expectation suggests a general factor may occur (p. 192, italics in original). Gorsuch (1983) recommends that if a general factor is perceived to exist a more appropriate method is to obliquely rotate the data and extract higher order factors. Advantages of this procedure include that under certain conditions varimax solutions are relatively invariant and a number of computer programs exist for the application of varimax (Gorsuch, 1983).

Equimax is a logical extension of both quartimax and varimax rotational procedures. The "(e)quimax method attempts to achieve simple structure with respect to both the rows and columns of the factor loading matrix" (Dillon & Goldstein, 1985, p. 91). Therefore, this procedure attempts to have an equivalent number of variables loading on each factor. Gorsuch (1973) reports that the variance in a matrix is more evenly spread as one moves from quartimax to varimax and then to equimax rotations. One limitation of this procedure is its availability. A second limitation is that the varimax procedure more closely adheres to the SBIV theoretical tenets. Therefore, it is favored over both quartimax and equimax procedures.

Confirmatory Factor Analysis: An Introduction

"Confirmatory factor analysis tests hypotheses that a specified subtest of variables legitimately define a prespecified factor " (Gorsuch, 1983, p. 127). Unlike exploratory techniques, these hypotheses are formulated in advance. Similar to exploratory factor analysis, several different methods are available. Three methods are utilized here. These are Procrustes, multiple group and hierarchical factor analysis.

Procrustes

The Procrustes solution has been proposed as an alternative to exploratory methods such as simple structure. Gorsuch (1983) has summarized the difficulties surrounding exploratory mathematical models such as simple structure. He writes:

> The major problem is that the ultimate goal of simple structure, invariance, may not be met, as Guilford has demonstrated (Guilford, 1975, 1977; Guilford & Hoepfner, 1969, 1971; Guilford & Zimmerman, 1963). An additional limitation is that it is solely exploratory and does not include information that might already exist that would guide the rotation into a confirmation of a previously found solution, or at least into a solution more congruent with patterns of theorization in the area. Guilford's preferred solution is to establish a matrix of hypothesized factor loadings, and use that as a target towards which the observed data are rotated (p. 232).

Guilford's "preferred solution" is called a Procrustes solution after the mythic Greek Inn Keeper who would stretch or cut his patrons to fit his beds. This label was first applied by Hurley & Cattell (1962):

> It will be recalled that the Greek hero Theseus encountered in his wanderings a character called Procrustes, whose beds fitted all travelers. Those who were too short for his beds he cruelly stretched and those who were too tall he cut down to size. If an investigator is satisfied - as many are - to announce that the fit is good, from visual judgement, then this program lends itself to the brutal feat of making almost any data fit almost any hypothesis! Because of this proclivity we gave the ... (computer)... code name Procrustes to this program, for this reference describes what it does for better or worse (p. 260).

Hurley and Cattell (1962) also state that statistical significance tests of goodness of fit have provided an answer, although imperfect, to evaluate Procrustean solutions. Hurley and Cattell (1962) argue that much is needed to be done on refining this technique but that the advantage over visual techniques are clear. Reddon, Marceau, and Holden (1985) note that several subsequent studies have criticized the use of Procrustean techniques on the basis that "the possibility exists for a considerable amount of capitalization on chance or overfitting in confirmatory factor rotations." Horn (1967) examined a hypothesized factor matrix based upon seventy-four random variables and a sample of three hundred. He concluded that "if an investigator were willing to interpret relatively low loadings, if these seemed to make sense, he needn't both to gather actual data: random variables may be labeled arbitrarily and pushed into solutions that make quite good sense" (p. 820).

Humphreys, Ilgen, McGrath, and Montanelli (1969) extend Horn's (1967) study by varying the number of subjects, variables and factors. They also look at confirmatory as well as exploratory rotations. Like Horn (1967), they found "that seemingly meaningful rotated factors can be obtained from the intercorrelations of random normal deviates whether the rotations are of the hypothesis testing or hypothesis seeking variety (Humphreys, Ilgen, McGrath & Montanelli, 1969, pp. 267-268). These authors conclude that increased care is required to support claims that the obtained factors are nonrandom. The following research design guidelines are suggested: a) at least four marker variables per factor should be employed b) sample size should be as large as possible and c) that the least number of variables compatible with research goals should be used. Similar conclusions and recommendations are offered by Nesselroade and Baltes (1970), and Nesselroade, Baltes and Labouvie (1971).

Jackson and Morf (1974), and Heeler and Whipple (1976) illustrate procedures that can be used to examine "whether rotation of a factor matrix to an hypothesized target produces a more adequate solution than rotation to a randomly determined target" Jackson and Morf (1974, p. 303). Jackson and Morf (1974) advocate the comparison of goodness of fit indices for random-based and hypothesis-based target matrices. They maintain "If the actual structure underlying a correlation matrix is in no way related to the hypothesized structure, or if there is no structure at all, then rotation to an hypothesis-based target matrix should yield a solution similar in terms of indices of goodness of fit to a rotation to a randomly determined target matrix" (p. 304). They conclude that such a procedure be routine until various relationships between sample size, number of tests and factors is better known.

The oblique procrustes procedures used here is described by Hendrickson and White (1966). This procedure is theory as opposed to analysis driven. A specialized application of Hendrickson and White's (1966) software program is employed. The orthogonal procrustes techniques are those designed by Reddon (1990).

Gorsuch (1983) warns that Procrustes solutions often results in factor correlations that are "unacceptably high" (p. 233). He emphasizes that the absence of an exact significance test for this procedure warrants the use of alternative methods. Gorsuch (1983) specifically advocates the use of confirmatory maximum likelihood (also see Heeler & Whipple, 1976) and multiple group factor analytic techniques because they enable the significance testing of hypotheses relates to correlations among factors and specific elements.

Holden, Reddon, Jackson and Helmes (1983), however, psychometrically examined the constructs underlying the multi-scale Basic Personality Inventory (BPI) using a Procrustean analysis. A parallel analysis was also undertaken using three random sets of data with sample sizes equivalent to the real data sets. The purpose of this analysis was to examine the degree to which chance was involve in the confirmatory rotation. "(T)he analysis of random data sets was also undertaken to offer a baseline from which possible distortions attributable to measurement error might be evaluated (Holden, Reddon, Jackson & Helmes, 1983, p. 40). Their results support both the theoretical and item structure of the BPI. In addition, evidence was also obtained to indicate that statistical artifacts did not compromise the observed factor structure. Therefore, support for their analysis was obtained when specific procedures designed to examine the influence of chance factors and their relationship to hypothetical factor structures are employed. Therefore, although the debate in the literature exists, it is evident that under certain restrictions the Procrustes solution can contribute to the analysis of hypothetical structures in well defined studies. This is particularly true when supplementary analyses such as goodness of fit indices are employed.

Broadbooks and Elmore (1987) stress that there are many goodness of fit statistics available. Their focus, however, is on one particular similarity coefficient, the congruence coefficient. They write:

> The congruence coefficient (referred to as the unadjusted correlation in some earlier studies) was developed as a measure of the similiarity of the factor patterns for different samples of subjects (Burt, 1948; Tucker, 1951; Wrigley & Neuhaus, 1955). The congruence coefficient involves a comparison of two sets of factor loadings in terms of both the pattern and magnitude of the loadings. It has been used extensively as a descriptive statistic in research comparing factors across studies (Broadbooks & Elmore, 1987, pp. 1-2).

Like the correlation coefficient, the congruence coefficient has a range of +1 (complete agreement) to -1 (inverse agreement). Similar to the correlation coefficient a value of zero indicates no or a lack of agreement. Several studies have examined various aspects of the congruence coefficient and these are outlined by Broadbooks and Elmore (1987). The latter's research indicates that the stability of the congruence coefficient increases with the number of subjects and variables studied. They also conclude that when the congruence coefficient is greater than .50 it will usually be a conservative estimate of the actual population value. Broadbooks and Elmore's (1987) results are supportive of the congruence coefficient and its use in this study as a goodness of fit index.

The congruence coefficient is only one means of examining the goodness of fit between hypothesized and obtained results. A second method is randomization or permutation tests. These tests are a fairly recent in their common use in applied statistics because of their reliance upon computers for the large number of calculations required (Edgington, 1987; Maguire, 1986). Such tests are an excellent means for testing hypotheses about one's results. The hypothesis tested is - "that the observed results differ systematically from a random allocation of data..." (Maguire, 1986, p. 168). Edgington (1987) summarizes this procedure as follows:

> A statistic is computed for the experimental data, then the data are permuted (divided or rearranged) repeatedly in a manner consistent with random assignment procedure, and the test statistic is computed for each of the resulting permutations. These data permutations, including the one representing the obtained results, constitute the reference set for determining significance" (p. 1).

Randomization tests are employed as a subanalysis to the Procrustean analysis. This is because of their utility in examining the fit of predicted models. They are one aspect of the overall construct validation process.

Multiple-Group Analysis

Multiple group factor analysis and its variations are considered excellent approaches for determining the existence of single or multiple factor solutions (Nunnally, 1967; For detailed technical discussions of this procedure see Gorsuch, 1983; Harman, 1976; and McDonald, 1985). It is an accepted procedure for examining subscale items within a particular test. Multiple-group factor analysis is an especially valuable tool for investigating theoretical structures that are purported to exist within a test battery.

Nunnally (1976) states that when examining the results of a multiple group analysis "it is appropriate to ask whether or not it is possible to find *disconfirming* results regarding the existence of factors" (p. 346). Three possible types of disconfirming information are identified. First, when low or negative correlations are found between variables (e.g., tests) and factors that are theorized to be related. Secondly, when variables hypothesized to be related to one factor correlate equally or more highly with another factor. A final type of disconfirming information identified by Nunnally (1967) is obtained when the absolute values in the residual matrix remain high after all hypothesized factors have been extracted. Such a finding suggests that the hypothetical factor structure does not account for a substantial amount of common variance.

A factor in most multiple group analyses is obtained by summing scores from specific variables that are hypothesized to define that factor. Typically each variable is assigned equal weight. These original variable scores are correlated with the multiple-group factor scores to obtain the factor structure (Gorsuch, 1983).

Multiple-group analyses can involve correlated or uncorrelated solutions. Correlated solutions are more common and preferred. This study follows this established tradition and restricts its analysis to oblique solutions. Gorsuch (1983) has summarized the difficulties inherent in orthogonal multiple-group analysis:

> The basic question is that of which factor gets the overlapping variance. If factor A is extracted first, the overlapping variance will be given to A; if factor B is extracted first, then B will get the overlap. If uncorrelated factors are to be extracted, the theory must specify the order of extraction so that hypotheses are given a fair test (p. 88).

The multiple group procedures employed are those documented by Paunonen (1987). Solutions derived by oblique multiple group most closely parallel those used by Thorndike, Hagen and Sattler (1986b). The multiple group method allows for solutions that are readily interpretable from a psychological standpoint and are consistent with SBIV test practices (i.e., factors can be calculated as the unweighted sum of each scales standardized item responses). The software used is that provided by the author (Paunonen, 1987).

Hierarchical Analysis

As noted in the literature review, during the 1930s a pivotal debate existed "between advocates of oblique simple structure and advocates of the notion of a general factor (i.e., one on which all the variables have nonzero regressions)... (McDonald, 1985, p. 104). Spearman saw oblique simple structure as a methodological artifact designed to avoid a general factor. Others saw *g* as a means of explaining the correlations that were found to exist between Thurstone's primary abilities (McDonald, 1985). Several authors including McNemar (1964), Gorsuch (1983) and McDonald (1985) have reconciled the diverse nature of findings by examining the scope of various intelligence measures and higher order factoring. Gorsuch (1983) has summarized the relationship between a second aspect of generalizability and higher order factoring. He writes:

> Primary factors indicate areas of generalizability. More generalization can occur within a factor than across factors, but this does not eliminate generalization across factors. When factors are correlated, some generalization is possible. These areas of generalization across the primary factors form the higher-order factors.
>
> The essential difference between the primary factors and the higher-order factors is that the primary factors are concerned with narrow areas of generalization. In some analyses, the reduction in accuracy when going from primary to second-

> order factors will be small; in other studies it may be quite great. It depends on the data being analyzed (p. 240).

The matrix of correlations among primary or first order factors, if they are rotated obliquely, can be factored in the same way as the original correlation matrix. A number of extraction procedures are available and these are discussed in detail by Gorsuch (1983) and McDonald (1985). The extraction and factoring process can be repeated if necessary. It is typically stopped when a single factor or uncorrelated factors result (Gorsuch, 1983).

Gorsuch (1983) notes that critics often suggest that higher-order factor are too far removed from the real world. He states that this criticism is unfounded as such factors are measured with the same degree of accuracy as the original factor. To ensure that one is not stepping too far from reality, the relationship between each level of higher-order factors and the original variables is often determined in addition to relationships between levels of factors. The presence of a strong theoretical framework for one's data is also viewed as an asset for understanding and identifying factors (Gorsuch, 1983).

Gorsuch (1983) correctly observed that oblique rotations, by their very nature, suggest overlapping factors are believed to exist. Such areas of overlap are said to be potentially indicative of higher order factors. The extraction of these higher order factors is often undertaken to gain a more complete understanding of one's data.

Several hierarchical arrangements are examined using the procedures described by Hendrickson and White (1966). These are primarily selected on the basis of the SBIV theoretical model , logical variants and the research findings previously reported. The solutions examined are believed to be an adequate representation of potential alternative structures offering a comprehensive evaluation of SBIV theoretical structure. It is to this evaluation we now turn.

CHAPTER 4

RESULTS

Introduction

The results of the various analysis are presented in this chapter. Descriptive statistics for each subsample are presented and compared to their standardization sample counterparts. Findings from the factor analytic procedures outlined in Chapter 3 are displayed. First, the unrotated principal components are identified. Next, they are rotated according to varimax orthogonal procedures. Orthogonal Procrustean methods are then employed. An examination of subtest scatter is then undertaken. This is followed by a multiple group analyses. Hierarchical factor analytic techniques are the final method used to examine the SBIV.

Descriptive Statistics

Table 4.0 presents descriptive scores for the minimum-maximum values, medians, means and standard deviations for each of the three subsamples and the entire sample. Product-moment intercorrelation coefficients between tests and Area Scores for each of these groups are also displayed. These tables can be compared to the standardization data presented in tables 4.2 to 4.5 and 6.1 in the *Technical Manual'*. The restricted presentation of standardization sample data prohibits direct comparisons. Therefore, this discussion will focus on theoretical as well as weighted standardization sample data.

Descriptive Statistics for the 3 to 6 Years Subsample. Table 4.0 provides the total number of cases, means and standard deviations for the youngest subsample by subtest. These results are provided primarily for information purposes as several of the subtests are later excluded because of their low rate of administration. These results differ slightly from those in Table 4.1 due to rounding procedures. Area and Composite Scores differ due to actual differences because of the number of subtests included in the analysis.

TABLE 4.0

SBIV Number of Cases, Means and Standard Deviations

for Subsample 3 to 6 Years by Subtest, Area and Composite Scores

Variable	N	Mean	SD
Vocabulary	130	55.6	7.0
Comprehension	130	55.4	6.6
Absurdities	130	55.4	6.6
Verbal Relations	0	.	.
Pattern Analysis	130	54.6	9.2
Copying	127	52.1	7.6
Matrices	12	56.3	3.2
Paper Folding & Cutting	0	.	.
Quantitative	130	54.4	7.8
Number Series	5	55.6	3.8
Bead Memory	130	51.2	8.5
Equation Building	0	.	.
Memory For Sentences	130	52.9	8.0
Memory For Digits	46	55.9	5.8
Memory For Objects	38	57.4	6.4
Verbal Reasoning	130	112.7	12.6
Abstract/Visual Reasoning	130	108.2	15.9
Quantitative Reasoning	130	108.2	16.0
Short-Term Memory	130	106.7	15.8
Composite Score	130	110.8	13.8

TABLE 4.1

Descriptive Statistics for the 3 to 6 Years Subsample

Test	Minimum	Maximum	Median	Mean	Standard Deviation
Vocabulary	37.0	76.0	56.0	55.6	7.2
Comprehension	37.0	75.0	55.0	55.3	6.9
Pattern Analysis	37.0	81.0	56.0	56.7	8.7
Copying	34.0	79.0	53.0	53.3	7.7
Quantitative	34.0	69.0	57.0	56.1	7.6
Bead Memory	32.0	79.0	53.0	51.7	9.1
Memory for Sentences	30.0	79.0	51.0	52.3	8.8
Reasoning Scores					
Verbal	81.0	148.0	113.0	112.7	12.9
Abstract/Visual	70.0	162.0	111.0	111.8	15.4
Quantitative	63.0	138.0	114.0	111.4	16.1
Short-Term Memory	64.0	152.0	107.0	106.8	17.2
Composite	77.0	158.0	114.0	112.9	14.1

Table 4.1 outlines minimum-maximum values, medians, means and standard deviations for the subtests retained for the 3 to 6 years subsample. this is the actual data used in the subsequent factor analyses. Standard Age Test Scores are wide ranging. A comparison of medians and means for each test suggest that outliers are not of concern. Test means range from a low of 51.2 (Bead Memory) to a high of 55.6 (Vocabulary). These are well within the empirical and theoretical value of 50 set by Thorndike, Hagen and Sattler (1986(a)(b)). Test standard deviations range from 6.7 (Comprehension) to 9.2 (Pattern Analysis). The majority fall within a single value of that proposed by the test authors.

Data for the Area scores for this subsample is, of course, parallel. Area means range from a low of 106.7 (Short-Term Memory) to a high of 112.7 (Vocabulary). The Composite Score of 110.8 is substantially higher than that proposed by Thorndike, Hagen and Sattler (1986(a)(b)). These results are, however, consistent with Vernon's (1987) suggestion that Canadians should score higher than their American counterparts because of their proportionately smaller minority group population. Approximately 13% of the Clinic sample is a member of an identifiable minority group. This in comparison to approximately 25% of the standardization sample. Other explanations such as higher education standards cannot, however, be ruled out.

Area score standard deviations for the youngest subsample range from 12.6 (Verbal Reasoning) to 16.0 (Quantitative). Three of the five Area Score standard deviations are within 2 points of the theoretical values proposed by the test authors. The Composite Score of 112.9 is consistent with overall elevated Area Scores. Aside from the generally higher values these results are consistent with those outlined in the *Technical Manual*'.

Transforming raw data matrices to either correlation or covariance matrices is customary prior to conducting a principal components analysis. The most common approach is to use correlation matrices based upon ordinary Pearson product-moment coefficients. Correlation coefficients are used to address the fact that frequently one is dealing with variables that have different units and scales (Dillon & Goldstein, 1985; for detailed discussions on the advantages and disadvantages of various correlational techniques as well as differences see Gorsuch, 1983; Nunnally, 1970; Rummel, 1970). "In computing a correlation coefficient between two variables, differences due to both the mean and the dispersion of the variables are removed. Thus, the transformation makes the variables directly comparable" (Dillon & Goldstein, 1985: p. 26).

Prior to discussing the correlational and other procedures a brief diversion is necessary to highlight the impact the nature of the sample can have on the obtained findings. Age and ability ranges are primary considerations. Comrey (1978) stresses that if ability tests are given to children who range widely in age, a general factor may emerge simply because most older children do better than younger children. As a result the "correlations between all tests will be very high, and a single factor will account for most of the variance. An unwary investigator might call this a general factor, but in reality it is only a maturation factor" (p. 650). This problem is lessened when age-related samples are employed.

A second consideration that is highlighted early by Comrey (1978) is that populations that are very selective reduce the variance on general factors, thereby highlighting group factors. As briefly noted in Chapter 3, the inclusion of subjects of all ability levels in each subsample is believed to offset, to some degree, the special nature of the Clinic population. Nevertheless, the fact that it is derived from a Clinic population should not be ignored. An additional concern identified by Comrey (1978) is the problem of outliers or unusual scores that can significantly influence correlation matrices. Outliers are potential sources of spurious common factor variance that can influence data interpretation. As discussed earlier, outliers are not believed to be an important variable in this study. Comrey, however, (1978) cautions the researcher to inspect all scores that are out of the ordinary for possible explanations and exclusion if necessary (Carroll, 1985). Several other issues of direct interest to prospective factor analytic investigators are raised in this article. In addition, he reviews a number of textbooks that are often referenced in factor analytic works. Several of these issues are discussed as they become relevant to data interpretations.

Comparison of the obtained correlation results to the standardization sample data is hampered by the fact that the *Technical Manual'* does not present these results according to the age groups utilized in the factor analysis. It is also important to reiterate that Thorndike, Hagen and Sattler (1986b) use median correlations rather than mean correlations to offset possible influences of extreme scores. The problems with this procedure are briefly noted in Chapter 3.

Table 4.2 displays the average intercorrelations for the 3 to 6 year subsample. The majority of correlations are positive and significant. Higher correlations are, of course, found between tests and their respective Area Scores (note intercorrelations of tests with Area and Composite Scores are not adjusted for overlap because the adaptive testing format of the SBIV does not require each subject to take all of the same tests). Exceptions to expected values are often the result of extremely low numbers of individuals taking one or both of the tests. Overall, the correlational results for the youngest subsample are generally consistent with those outlined in the *Technical Manual'*.

Descriptive Statistics for the 7 to 11 Years Subsample. Table 4.3 provides the total number of cases, means and standard deviations for the middle subsample by subtest. These results are provided primarily for information purposes as several of the subtests are excluded later because of their low administration rate. These results differe slightly from those in Table 4.4 due to rounding procedures. Area and Composite Score results differ due to actual differences in the number and type of tests included in the analysis. Table 4.4 outlines minimum-maximum values, medians, means and standard deviations for the 7 to 11 years subsample. This is the actual data used in the factor analysis. Similar to the youngest subsample Standard Age Scores for the test are wide ranging. A comparison of test medians and means suggests that outliers are not problematic. Test means range from a

TABLE 4.2

SBIV Intercorrelations Subsample for the 3 to 6 Years

	Verbal Reasoning Area				Abstract/visual Reasoning Area				Quantitative Reasoning Area			Short-Term Memory Area				Area and Composite Scores				
	Vocabulary	Comprehension	Absurdities	[Verbal Relations]	Pattern Analysis	Copying	[Matrices]	Paper Folding & Cutting	Quantitative	[Number Series]	Equation Building	Bead Memory	[Memory for Sentences]	[Memory for Digits]	Memory for Objects	Verbal Reasoning	Abstract/visual Reasoning	Quantitative Reasoning	Short-Term Memory	Composite
Vocabulary		.62**	.42**	.	.31**	.20**	-.20	.	.32**	.62	.	.32**	.54**	.59**	.18	.84**	.36**	.38**	.50**	.63**
Comprehension			.47**	.	.26*	.19	-.16	.	.28**	.35	.	.41**	.50**	.45**	.02	.86**	.26*	.36**	.50**	.59**
Absurdities				.	.25*	.25*	-.32	.	.26	-.08	.	.27*	.37**	.47**	.36	.73**	.28**	.27**	.42**	.53**
[Verbal Relations]					.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Pattern Analysis						.44**	-.22	.	.51**	.35**	.	.42**	.32**	.40*	.42*	.34**	.87**	.54**	.46**	.71
Copying							.36**	.	-.11	.33	.	.36**	.31**	.35	.41*	.29**	.80**	.42**	.42**	.61**
[Matrices]								.	-.17	1.0**	.	.13	.32	.15	-.05	-.25	-.01	-.19	.12	-.11
Paper Folding & Cutting									.	.	.	.	.	.	.	.	.	.	.	.
Quantitative										.	.	.48**	.36**	.42*	.35	.35**	.50**	.94**	.46**	.75**
[Number Series]											.	.48	-.46	.04	-.27	.01	.66	.97*	-.05	.65
Equation Building												.	.	.	.	.	.	.	.	.

* .01 tailed significance . if coefficient cannot be computed

** .001 tailed significance [ ] fails to meet the criteria of at least 65% of sample responding

TABLE 4.2 continued

SBIV Intercorrelations Subsample for the 3 to 6 Years

	Verbal Reasoning Area				Abstract/visual Reasoning Area				Quantitative Reasoning Area			Short-Term Memory Area				Area and Composite Scores				
	Vocabulary	Comprehension	Absurdities	[Verbal Relations]	Pattern Analysis	Copying	[Matrices]	Paper Folding & Cutting	Quantitative	[Number Series]	Equation Building	Bead Memory	[Memory for Sentences]	[Memory for Digits]	Memory for Objects	Verbal Reasoning	Abstract/visual Reasoning	Quantitative Reasoning	Short-Term Memory	Composite
Bead Memory													.45**	.35*	.32	.40**	.45**	.52**	.82**	.70**
Memory for Sentences														.62**	.44*	.58**	.39**	.43**	.81**	.69**
[Memory for Digits]															.40*	.60**	.45**	.42	.71**	.64**
[Memory for Objects]																.23	.44*	.34	.67**	.52**
Verbal Reasoning																	.37**	.41**	.58**	.71**
Abstract/visual Reasoning																		.56**	.52**	.78**
Quantitative Reasoning																			.53**	.79**
Short-Term Memory																				.83**
Composite																				

* .01 tailed significance   . if coefficient cannot be computed

** .001 tailed significance   [ ] fails to meet the criteria of at least 65% of sample responding

TABLE 4.3

SBIV Number of Cases, Means and Standard Deviations

for Subsample 7 to 11 Years by Subtest, Area and Composite Scores

Variable	N	Mean	SD
Vocabulary	147	50.7	9.1
Comprehension	147	52.2	8.8
Absurdities	144	51.5	8.5
Verbal Relations	1	66.0	.
Pattern Analysis	147	51.3	8.6
Copying	129	46.3	8.8
Matrices	87	51.2	7.9
Paper Folding & Cutting	2	56.5	3.5
Quantitative	147	48.6	7.5
Number Series	80	52.3	8.6
Equation Building	2	60.5	10.6
Bead Memory	147	49.1	10.1
Memory For Sentences	147	48.8	9.6
Memory For Digits	124	50.1	7.8
Memory For Objects	117	51.4	7.8
Verbal Reasoning	147	103.4	17.7
Abstract/Visual Reasoning	147	99.0	17.0
Quantitative Reasoning	147	99.1	15.6
Short-Term Memory	147	99.7	19.0
Composite Score	147	100.3	18.1

TABLE 4.4

Descriptive Statistics for the 7 to 11 Years Subsample

Test	Minimum	Maximum	Median	Mean	Standard Deviation
Vocabulary	31.0	71.0	50.0	50.7	9.1
Comprehension	23.0	73.0	52.0	52.2	8.8
Absurdities	32.0	70.0	52.0	51.6	8.5
Pattern Analysis	23.0	69.0	52.0	51.3	8.6
Copying	23.0	67.0	47.0	46.4	8.4
Quantitative	25.0	74.0	48.0	48.6	7.5
Bead Memory	25.0	72.0	49.0	49.2	10.1
Memory for Sentences	27.0	75.0	47.0	48.8	9.6
Memory for Digits	32.0	79.0	50.0	50.1	7.4
Memory for Objects	33.0	70.0	51.0	51.4	7.2
Reasoning Scores					
Verbal	60.0	142.0	105.0	103.4	17.7
Abstract/Visual	43.0	131.0	101.0	99.0	17.0
Quantitative	50.0	148.0	100.0	99.2	15.6
Short-Term Memory	51.0	150.0	99.0	99.8	19.0
Composite	49.0	143.0	102.0	100.3	18.1

low of 46.4 (Copy) to 52.2 (Comprehension). Standard deviations are in the range of 7.2 (Memory for Objects) to 10.1 (Bead Memory).

Area Score data, also suggests that the middle group results, relative to the youngest group, more closely parallel those proposed to exist by Thorndike, Hagen and Sattler (1986(a)(b)). Four of five Area Scores are within a single point of the hypothesized means. Like the youngest subsample the Verbal Reasoning Score is higher relative to the other Area Scores.

Table 4.5 outlines the average intercorrelations for the 7 to 11 year subsample. Similar to the younger group the intercorrelations are typically positive and statistically significant. Relative to the younger group, however, the correlations are generally larger. A similar pattern is observed in the standardization data. Correlations between tests and related Area Scores are in the moderately high range and follow predicted patterns. Correlations between Area and Composite Scores are generally at the .90 level.

Descriptive Statistics for the 12 to 23 Years Subsample. Table 4.6 provides the total number of cases, means and standard deviations for the oldest subsample by subtest. Then results are provided primarily for information purposes as several of the subtests are excluded later because of their low administration rate. These results differe slightly from those in Table 4.7 due to rounding procedures. Area and Composite Score results differ due to actual differences in the number and type of tests included in the analysis. Table 4.7 outlines minimum-maximum values, medians, means and standard deviations for the 12 to 23 years subsample. This is the data actually used in the subsequent factor analyses. Like the two younger groups, Standard Age Scores are wide ranging. A comparison of median and mean values suggests that outliers are not a major concern. Test means range from a low of 44.5 (Memory for Sentences) to 50.1 (Pattern Analysis). The Standard deviations for these scores are in the range from 8.0 (Absurdities) to 11.1 (Bead Memory).

TABLE 4.5

SBIV Intercorrelations Subsample for the 7 to 11 Years

	Verbal Reasoning Area				Abstract/visual Reasoning Area				Quantitative Reasoning Area			Short-Term Memory Area				Area and Composite Scores				
	Vocabulary	Comprehension	Absurdities	[Verbal Relations]	Pattern Analysis	Copying	[Matrices]	Paper Folding & Cutting	Quantitative	[Number Series]	[Equation Building]	Bead Memory	[Memory for Sentences]	[Memory for Digits]	Memory for Objects	Verbal Reasoning	Abstract/visual Reasoning	Quantitative Reasoning	Short-Term Memory	Composite
Vocabulary		** .75	** .64	.	** .52	** .37	** .67	1.0	** .58	** .72	1.0	** .49	** .69	** .59	** .56	** .91	** .61	** .67	** .67	** .80
Comprehension			** .58	.	** .59	** .37	** .61	1.0	** .65	.66	1.0	** .50	** .63	** .55	** .55	** .89	** .63	** .70	** .65	** .80
Absurdities				.	** .58	** .40	** .71	.	** .48	.69	.	** .41	** .50	** .41	** .57	** .84	** .64	** .58	** .62	** .74
[Verbal Relations]					.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Pattern Analysis						** .49	** .59	.	** .59	** .62	1.0	** .66	** .55	** .51	** .58	** .63	** .87	** .64	** .71	.80
Copying							* .35	.	** .45	** .46	.	** .38	** .37	** .42	** .51	.43	.81	.53	.50	.63
[Matrices]								1.0	** .54	** .71	.	** .54	** .59	** .43	** .66	** .74	** .79	** .68	** .68	** .80
[Paper Folding & Cutting]									1.0	1.0	.	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Quantitative										.63	1.0	** .48	** .57	** .58	** .54	** .65	** .64	** .93	** .64	.79
[Number Series]											1.0	** .54	** .57	** .59	** .56	** .77	** .74	** .91	** .66	** .86
[Equation Building]												1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

* .01 tailed significance

** .001 tailed significance

. if coefficient cannot be computed

[ ] fails to meet the criteria of at least 65% of sample responding

TABLE 4.5 continued

SBIV Intercorrelations Subsample for the 7 to 11 Years

	Verbal Reasoning Area				Abstract/visual Reasoning Area				Quantitative Reasoning Area			Short-Term Memory Area				Area and Composite Scores				
	Vocabulary	Comprehension	Absurdities	[Verbal Relations]	Pattern Analysis	Copying	[Matrices]	Paper Folding & Cutting	Quantitative	[Number Series]	[Equation Building]	Bead Memory	Memory for Sentences	Memory for Digits	Memory for Objects	Verbal Reasoning	Abstract/visual Reasoning	Quantitative Reasoning	Short-Term Memory	Composite
Bead Memory													.53**	.48**	.52**	.57**	.64**	.53**	.82**	.73**
Memory for Sentences														.70**	.55**	.69**	.59**	.62**	.83**	.78
Memory for Digits															.52**	.59**	.55**	.63**	.78**	.72
Memory for Objects																.62**	.67**	.60**	.78**	.74**
Verbal Reasoning																	.71**	.74**	.74**	.89**
Abstract/visual Reasoning																		.73**	.76**	.89**
Quantitative Reasoning																			.69**	.88**
Short-Term Memory																				.90**
Composite																				1.0

* .01 tailed significance　　. if coefficient cannot be computed

** .001 tailed significance　　[ ] fails to meet the criteria of at least 65% of sample responding

TABLE 4.6

SBIV Number of Cases, Means and Standard Deviations

for Subsample 12 to 23 Years by Subtest, Area and Composite Scores

Variable	N	Mean	SD
Vocabulary	94	49.3	9.9
Comprehension	94	48.5	10.5
Absurdities	73	46.8	8.5
Verbal Relations	39	55.1	7.9
Pattern Analysis	94	50.1	8.6
Copying	54	40.2	8.2
Matrices	80	50.1	10.0
Paper Folding & Cutting	43	55.5	7.5
Quantitative	94	47.7	11.0
Number Series	71	50.7	9.2
Equation Building	27	62.5	9.4
Bead Memory	94	47.0	11.1
Memory For Sentences	94	44.5	9.9
Memory For Digits	88	47.7	8.7
Memory For Objects	86	47.5	8.3
Verbal Reasoning	94	97.5	20.4
Abstract/Visual Reasoning	94	97.6	21.9
Quantitative Reasoning	94	97.9	23.8
Short-Term Memory	94	91.4	21.1
Composite Score	94	95.4	23.4

TABLE 4.7

Descriptive Statistics for the 12 to 23 Years Subsample

Test	Minimum	Maximum	Median	Mean	Standard Deviation
Vocabulary	26.0	69.0	50.0	49.3	9.9
Quantitative	26.0	74.0	46.0	47.7	11.0
Bead Memory	23.0	71.0	47.0	47.0	11.1
Memory for Sentences	26.0	68.0	46.0	44.5	9.9
Pattern Analysis	32.0	63.0	54.0	50.1	8.6
Comprehension	24.0	75.0	48.0	48.5	10.5
Absurdities	32.0	66.0	49.0	48.0	8.0
Matrices	29.0	68.0	50.0	49.0	10.1
Number Series	31.0	68.0	51.5	49.5	9.8
Memory for Digits	31.0	71.0	46.0	48.0	8.9
Memory for Objects	32.0	70.0	48.0	47.6	8.1
Reasoning Scores					
Verbal	48.0	144.0	99.5	97.5	20.4
Abstract/Visual	52.0	135.0	102.5	97.6	21.9
Quantitative	52.0	151.0	97.0	97.9	23.8
Short-Term Memory	48.0	144.0	89.0	91.4	21.1
Composite	47.0	139.0	99.0	95.4	23.4

Area Score means for the oldest subsample range from 91.4 (Short-Term Memory) to 97.9 (Quantitative Reasoning). A downward trend of Area scores means from youngest to oldest subsamples is observed. A more direct relationship is found between subsample age and variability.

Table 4.8 provides the average intercorrelations for the 12 to 23 year subsample. All correlations are positive and the majority are highly significant. Relative to the two younger groups the correlations are larger. Test and Area Score intercorrelations are moderately high and in the predicted direction. Correlations between Area and Composite Scores are very high ranging from .91 (Abstract/visual Reasoning) to .95 (Verbal Reasoning).

Descriptive Statistics for the 3 to 23 Year Sample. The minimum-maximum values, medians, means and standard deviations for the entire sample by subtest are displayed in Table 4.9. The number of valid and cases are provided to make explicit data that is based on few subjects. Data is presented for all tests even though many are excluded from the primary analysis.

A review of median and mean scores indicates that they are remarkably consistent. Test means range from a low of 47.6 (Copy) to 62.3 (Equation Building). The latter figure is based on the fewest number of protocols and is, therefore, suspect. Overall, test means based on acceptable proportions of the sample are in line with hypothesized values. Test standard deviations are consistent with those set forth by Thorndike, Hagen and Sattler (1986(a)(b)) and range from 7.3 (Paper Folding and Cutting) to 10.0 (Bead Memory).

Area means are close to those proposed by the test authors. They range from 101.9 (Abstract/visual Reasoning) to 105.1 (Verbal Reasoning). The Composite Score of 102.8 is consistent but not necessarily supportive of Vernon's (1987) hypothesis that Canadian samples should score higher than American samples. There is some evidence of greater variability among the Clinical sample than the weighted standardization sample as Area standard deviations range from 17.9 (Verbal) to 19.4 (Short-term Memory).

TABLE 4.8

SBIV Intercorrelations Subsample for the 12 to 23 Years

	Verbal Reasoning Area				Abstract/visual Reasoning Area				Quantitative Reasoning Area			Short-Term Memory Area				Area and Composite Scores				
	Vocabulary	Comprehension	Absurdities	[Verbal Relations]	Pattern Analysis	[Copying]	Matrices	[Paper Folding & Cutting]	Quantitative	Number Series	[Equation Building]	Bead Memory	Memory for Sentences	Memory for Digits	Memory for Objects	Verbal Reasoning	Abstract/visual Reasoning	Quantitative Reasoning	Short-Term Memory	Composite
Vocabulary		** .80	** .75	** .62	** .68	** .69	** .73	** .50	** .72	** .76	** .58	** .69	** .77	** .61	** .58	** .93	** .76	** .77	** .80	** .87
Comprehension			** .77	** .58	** .66	** .64	** .73	** .69	** .80	** .80	** .52	** .68	** .75	** .60	** .55	** .93	** .74	** .83	** .78	** .88
Absurdities				** .65	** .75	** .74	** .73	** .64	** .67	** .70	* .58	** .64	** .69	** .59	** .55	** .89	** .80	** .72	** .73	** .85
[Verbal Relations]					.33	.34	** .53	** .65	** .55	** .79	** .62	.36	* .42	.37	* .39	** .77	** .49	** .68	** .50	** .68
Pattern Analysis						** .73	** .72	** .52	** .68	** .74	.32	** .69	** .60	** .59	** .52	** .75	** .86	** .73	** .71	** .82
[Copying]							** .60	.38	** .59	** .66	** .30	** .66	** .58	** .58	** .55	** .74	** .86	** .63	** .68	** .76
Matrices								** .75	** .74	** .78	* .51	** .71	** .70	** .66	** .51	** .80	** .89	** .80	** .78	** .88
[Paper Folding & Cutting]									** .71	** .75	** .72	** .49	** .54	* .39	.29	** .69	** .85	** .79	** .56	** .80
Quantitative										** .83	** .70	** .66	** .72	** .67	** .49	** .81	** .75	** .97	** .77	** .89
Number Series											** .67	** .69	** .71	** .62	** .52	** .86	** .81	** .94	** .79	** .93

* .01 tailed significance
** .001 tailed significance
. if coefficient cannot be computed
[ ] fails to meet the criteria of at least 65% of sample responding

TABLE 4.8 continued

SBIV Intercorrelations Subsample for the 12 to 23 Years

	Verbal Reasoning Area				Abstract/visual Reasoning Area				Quantitative Reasoning Area			Short-Term Memory Area				Area and Composite Scores				
	Vocabulary	Comprehension	Absurdities	[Verbal Relations]	Pattern Analysis	[Copying]	Matrices	[Paper Folding & Cutting]	Quantitative	Number Series	[Equation Building]	Bead Memory	Memory for Sentences	Memory for Digits	Memory for Objects	Verbal Reasoning	Abstract/visual Reasoning	Quantitative Reasoning	Short-Term Memory	Composite
[Equation Building]												.29	.29	.34	.25	** .53	** .65	** .90	.40	** .77
Bead Memory													** .62	** .64	** .53	** .74	** .74	** .73	** .86	** .82
Memory for Sentences														** .68	** .57	** .81	** .71	** .75	** .87	** .83
Memory for Digits															** .59	** .67	** .68	** .71	** .87	** .78
Memory for Objects																** .62	** .56	** .55	** .79	** .67
Verbal Reasoning																	** .83	** .86	** .85	** .95
Abstract/visual Reasoning																		** .81	** .80	** .91
Quantitative Reasoning																			** .82	** .94
Short-Term Memory																				** .92
Composite																				

* .01 tailed significance
** .001 tailed significance
. if coefficient cannot be computed
[ ] fails to meet the criteria of at least 65% of sample responding

TABLE 4.9

Descriptive Statistics for the 3 to 23 Years Sample

Test	Minimum	Maximum	Median	Mean	Standard Deviation	Valid Cases	Missing Cases
Vocabulary	26.0	76.0	53.0	52.0	9.0	371	0
Comprehension	23.0	75.0	53.0	52.4	9.0	371	0
Absurdities	32.0	73.0	52.0	52.0	8.5	347	24
Verbal Relations	36.0	72.0	56.5	55.4	8.0	40	331
Pattern Analysis	23.0	81.0	53.0	52.2	9.0	371	0
Copy	23.0	79.0	48.0	47.6	9.3	310	61
Matrices	30.0	68.0	52.0	51.1	8.8	179	192
Paper Folding & Cutting	40.0	67.0	55.0	55.6	7.3	45	326
Quantitative	25.0	74.0	50.0	50.4	9.1	371	0
Number Series	31.0	75.0	52.0	51.7	8.8	156	215
Equation Building	46.0	82.0	62.0	62.3	9.3	29	342
Bead Memory	23.0	79.0	50.0	49.3	10.1	371	0
Memory for Sentences	26.0	79.0	49.0	49.1	9.6	371	0
Memory for Digits	31.0	79.0	50.0	50.3	8.3	258	113
Memory for Objects	32.0	72.0	51.0	51.0	8.4	241	130
Reasoning Areas							
Verbal	48.0	148.0	107.0	105.1	17.9	371	0
Abstract/Visual	43.0	162.0	103.0	101.9	18.6	371	0
Quantitative	50.0	151.0	102.0	102.1	18.7	371	0
Short-Term Memory	48.0	152.0	102.0	100.1	19.4	371	0
Composite Score	47.0	158.0	105.0	102.8	19.3	371	0

Table 4.10 outlines the intercorrelations for the combined Clinic sample. It can be compared to Table 6.1 of the *Technical Manual'*. Generally, the Clinic sample intercorrelations are remarkably consistent with the standardization sample. The Clinic sample intercorrelations are, however, somewhat higher. Exceptions to this pattern are typically confined to tests that reflect the fewest subjects. Overall, the descriptive statistics for all three subsamples and the entire sample are compatible with the SBIV standardization data. They provide initial support for subsequent analyses.

Scatter Analysis

Correlations of scatter indices (range of scaled scores across core tests and Areas) with each of the Area scores, the Composite score and two demographic variables are displayed in Table 4.11. The correlations between the demographic variables and scatter are small (-.02 to -.09). Smaller correlations are also obtained between Short-Term Memory Reasoning and the two scatter indices (.03 and .04). Slightly higher correlations are found between Abstract/visual Reasoning and the Area Score range (.11), Quantitative Reasoning and both the Core tests (.11) and Area Scores (.12). Significanct but small correaltions are found between both scatter indices and the Composite (.15 and .14) as well as Verbal Reasoning Scores (.21 and .26). The Abstract/visual Reasoning Score is also significantly correlated with the Core test range (.21) and Area score range (.26).

The above findings are supportive of examining scatter of various Composite score (IQ) levels. This is supported by Matarazzo, Daniel, Prifitera and Herman's (1988) work. Although the correlational analyses is not supportive of age related classifications, subsequent analyses and consistency of presentation provides such a rationale.

Tables 4.12 to 4.15 outline the percentage of cases at or above each level of overall (highest and lowest) Area score scatter (range) by composite Score level and for the various age groupings. Unlike Matarazzo, Daniel, Prifitera and Herman (1988) study, the relationship between Composite scores and mean/median scatter is less clear. No obvious

TABLE 4.10

SBIV Intercorrelations Subsample for the 3 to 23 Years

	Verbal Reasoning Area				Abstract/visual Reasoning Area				Quantitative Reasoning Area			Short-Term Memory Area				Area and Composite Scores				
	Vocabulary	Comprehension	Absurdities	Verbal Relations	Pattern Analysis	Copying	Matrices	Paper Folding & Cutting	Quantitative	Number Series	Equation Building	Bead Memory	Memory for Sentences	Memory for Digits	Memory for Objects	Verbal Reasoning	Abstract/visual Reasoning	Quantitative Reasoning	Short-Term Memory	Composite
Vocabulary		** .76	** .66	** .64	** .52	** .50	** .70	** .50	** .60	** .74	* .56	** .53	** .70	** .64	** .57	** .91	** .63	** .65	** .69	** .81
Comprehension			** .66	** .60	** .53	** .47	** .67	** .68	** .64	** .73	* .53	** .56	** .67	** .60	** .55	** .91	** .60	** .68	** .69	** .81
Absurdities				** .65	** .55	** .54	** .70	** .64	** .54	** .69	* .58	** .51	** .58	** .56	** .62	** .85	** .63	** .58	** .65	** .76
[Verbal Relations]					.35	.34	** .54	** .64	** .58	** .81	** .63	* .39	* .46	* .41	* .42	** .78	** .51	** .69	** .53	** .69
Pattern Analysis						.57	** .60	** .52	** .60	** .67	.36	** .60	** .52	** .54	** .56	** .60	** .86	** .64	** .64	** .77
Copying							** .48	.37	** .55	** .56	.30	** .48	** .50	** .56	** .61	** .56	** .85	** .59	** .60	** .72
[Matrices]								** .75	** .64	** .75	* .51	** .64	** .64	** .56	** .57	** .76	** .83	** .73	** .73	** .83
Paper Folding & Cutting									** .68	** .74	** .72	** .49	** .54	* .40	.29	** .68	** .85	** .77	** .55	** .78
[Quantitative]										** .74	** .68	** .55	** .60	** .63	** .53	** .67	** .67	** .95	** .66	** .84
[Number Series]											** .67	** .61	** .63	** .61	** .54	** .81	** .78	** .92	** .72	** .90

* .01 tailed significance
** .001 tailed significance
. if coefficient cannot be computed
[ ] fails to meet the criteria of at least 65% of sample responding

TABLE 4.10 continued

SBIV Intercorrelations Subsample for the 3 to 23 Years

	Verbal Reasoning Area				Abstract/visual Reasoning Area				Quantitative Reasoning Area			Short-Term Memory Area				Area and Composite Scores				
	Vocabulary	Comprehension	Absurdities	Verbal Relations	Pattern Analysis	Copying	Matrices	Paper Folding & Cutting	Quantitative	Number Series	Equation Building	Bead Memory	Memory for Sentences	Memory for Digits	Memory for Objects	Verbal Reasoning	Abstract/visual Reasoning	Quantitative Reasoning	Short-Term Memory	Composite
[Equation Building]												.31	.30	.36	.25	** .62	** .67	** .88	.39	** .75
Bead Memory													** .55	** .55	** .53	** .60	** .63	** .60	** .83	** .75
Memory for Sentences														** .71	** .59	** .73	** .60	** .63	** .85	** .79
[Memory for Digits]															** .59	** .67	** .62	** .65	** .83	** .77
[Memory for Objects]																** .64	** .62	** .57	** .81	** .72
Verbal Reasoning																	** .69	** .72	** .76	** .89
Abstract/visual Reasoning																		** .73	** .72	** .89
Quantitative Reasoning																			** .70	** .89
Short-Term Memory																				** .90
Composite																				

* .01 tailed significance
** .001 tailed significance
. if coefficient cannot be computed
[ ] fails to meet the criteria of at least 65% of sample responding

TABLE 4.11

Correlations Between Core Test and Area Score Scatter and Other Variables

Variables	N	Core Test Range	Area Score Range
Verbal Reasoning	371	.21**	.26**
Abstract/Visual Reasoning	371	.17**	.11
Quantitative Reasoning	371	.11	.12
Short-Term Memory Reasoning	371	.03	.04
Composite Score	371	.15*	.14*
Age	371	-.06	-.09
Sex	371	-.06	-.02

* if $p < .01$
** if $p < .001$

TABLE 4.12

SBIV: Percentage of Cases at or Above Each Level of Area Score Scatter by Composite Score Groupings for the 3 to 6 Years Subsample

Scatter Range	Composite Score Groupings						
	<65	66-82	83-99	100-116	117-133	134 >	ALL
56-60		0.0	0.0	1.7	0.0	0.0	0.8
51-55		0.0	0.0	1.7	2.6	0.0	1.5
46-50		0.0	0.0	3.4	2.6	0.0	2.3
41-45		25.0	4.2	6.9	5.1	0.0	5.4
36-40		25.0	4.2	8.6	7.7	0.0	7.7
31-25		25.0	20.8	20.7	20.5	0.0	20.0
26-30		50.0	33.3	34.5	43.6	0.0	36.2
21-25		50.0	75.0	52.0	53.8	40.0	56.2
16-20		50.0	87.5	70.7	79.5	80.0	75.4
11-15		75.0	95.8	89.7	94.9	100.0	92.3
6-10		100.0	100.0	91.4	97.4	100.0	98.5
0-5		100.0	100.0	100.0	100.0	100.0	100.0
N	0	4	24	58	39	5	130
$\bar{X}$	-	22.75	23.29	22.62	23.26	19.60	22.82
SD	-	12.84	8.11	11.08	9.93	4.72	10.0
MEDIAN	-	22.0	22.50	21.00	22.00	20.00	22.00

TABLE 4.13

SBIV: Percentage of Cases at or Above Each Level of Area Score Scatter

by Composite Score Groupings for the 7 to 11 Years Subsample

Scatter Range	Composite Score Groupings						
	<65	66-82	83-99	100-116	117-133	134 >	ALL
56-60	0.0	0.0	0.0	0.0	0.0	0.0	0.0
51-55	0.0	0.0	0.0	0.0	0.0	0.0	0.0
46-50	0.0	0.0	0.0	1.7	0.0	0.0	0.7
41-45	16.7	0.0	5.1	3.4	5.0	0.0	4.1
36-40	33.3	5.3	5.1	3.4	10.0	0.0	6.1
31-25	50.0	5.3	7.7	5.2	20.0	0.0	9.5
26-30	50.0	15.8	20.5	20.7	30.0	20.0	22.4
21-25	50.0	21.1	35.9	34.5	45.0	60.0	35.4
16-20	83.3	42.1	61.5	67.2	80.0	80.0	65.3
11-15	83.3	63.2	87.2	84.5	85.0	80.0	82.3
6-10	100.0	94.7	97.4	96.6	95.0	100.0	96.6
0-5	100.0	100.0	100.0	100.0	100.0	100.0	100.0
N	6	19	39	58	20	5	147
$\bar{X}$	25.50	15.21	18.97	18.79	21.35	19.20	19.01
SD	13.81	8.40	8.53	8.40	10.43	6.83	9.02
MEDIAN	25.50	13.00	17.00	18.00	18.00	20.00	18.00

TABLE 4.14

SBIV: Percentage of Cases at or Above Each Level of Area Score Scatter

By Composite Score Groupings for the 12 to 23 Years Subsample

Scatter Range	Composite Score Groupings						
	<65	66-82	83-99	100-116	117-133	134 >	ALL
56-60	0.0	0.0	0.0	0.0	0.0	0.0	0.0
51-55	0.0	0.0	0.0	0.0	0.0	0.0	0.0
46-50	0.0	5.3	0.0	0.0	5.9	0.0	2.1
41-45	0.0	5.3	0.0	0.0	17.6	0.0	4.3
36-40	0.0	5.3	0.0	7.7	17.6	0.0	6.4
31-25	0.0	5.3	0.0	19.2	23.5	0.0	10.6
26-30	21.4	21.1	13.3	30.8	41.2	0.0	25.5
21-25	21.4	36.8	33.3	42.3	62.5	66.6	40.4
16-20	42.9	57.9	73.3	69.2	62.5	66.6	61.7
11-15	73.3	68.4	100.0	76.9	88.2	100.0	81.9
6-10	100.0	100.0	100.0	92.3	94.1	100.0	96.8
0-5	100.0	100.0	100.0	100.0	100.0	100.0	100.0
N	14	19	15	26	17	3	94
$\bar{X}$	16.07	17.90	19.40	19.89	24.00	19.3	19.56
SD	7.10	10.10	4.64	9.77	12.90	7.23	9.57
MEDIAN	13.5	17.00	19.00	17.50	24.00	23.00	17.00

TABLE 4.15

SBIV: Percentage of Cases at or Above Each Level of Area Score Scatter by Composite Score Groupings for the 3 to 23 Years Subsample

Scatter Range	Composite Score Groupings						
	<65	66-82	83-99	100-116	117-133	134 >	ALL
56-60	0.0	0.0	0.0	0.7	0.0	0.0	0.3
51-55	0.0	0.0	0.0	0.7	1.3	0.0	0.5
46-50	0.0	2.4	0.0	2.1	2.6	0.0	1.6
41-45	5.0	2.4	3.8	4.2	7.9	0.0	4.6
36-40	10.0	7.1	3.8	6.3	10.5	0.0	6.7
31-25	15.0	7.1	10.3	14.1	21.1	0.0	13.5
26-30	25.0	21.4	23.1	28.2	39.5	7.7	28.0
21-25	25.0	31.0	47.4	43.0	52.6	46.2	43.9
16-20	55.0	50.0	70.5	69.0	75.0	76.9	67.9
11-15	80.0	66.7	92.3	85.2	90.8	92.3	85.7
6-10	100.0	97.6	98.7	96.5	96.1	100.0	97.3
0-5	100.0	100.0	100.0	100.0	100.0	100.0	100.0
N	20.	42	78	142	76	13	371
$\bar{X}$	18.90	17.14	20.39	20.56	22.92	19.39	20.49
SD	10.21	9.64	7.96	9.90	10.68	5.64	9.64
MEDIAN	16.0	15.50	19.50	18.00	21.50	20.00	19.00

patterns between these variables are evident for any of the four age groupings. Noteworthy is the fact that the lowest age group possessed the highest overall mean and median range despite the absence of any subjects in the lowest Composite score level.

These tables suggest that in this clinic sample a difference between an individual's highest and lowest Area score of 20 points is not unusual. This is true even when the subject is functioning at various levels of overall ability. Therefore, such a difference in and of itself is not necessarily a sign of pathology (Matarazzo, Daniel, Prifitera & Herman, 1988). Thus, although statistical significance (.05 level) between Areas in the standardization sample is reached anywhere from 6 to 18 points depending (see Technical Manual Table F.2, p. 134) upon Areas and ages that are compared, clinical significance, particularly in specialized populations, may be at a much higher level.

Table 4.16 is a pairwise comparison of the difference between Verbal and Abstract/visual area scores for each of the subsamples. Overall, the magnitude, mean, and median scatter indices between these two Area scores appear to decrease with age. Such a pattern is consistent with the standardization sample. However, the variability in this clinic sample is greater. Such a finding is not surprising given the somewhat specialized nature of clinic referrals.

Tables 4.17 and 4.18 provide the distribution of minimum and maximum Area score values. An obvious trend is evident. Average and median minimum and maximum values decrease with age. This pattern is evident across all three subsamples. This pattern may be unique to this Clinic population, artifactual (i.e., relatively higher floor for the younger age groups) or a valid finding.

In summation, the finding that consistent and readily identifiable Area Score scatter patterns are not found in this clinic population is not insignificant. Statistical artifacts such as high floors or low ceilings do not appear to be present (although for the youngest age group no examinees scored less than 65 but this may be a sampling as opposed to statistical artifact). Unlike Matarazzo, Daniel, Prifitera and Herman's (1988) an observable pattern

TABLE 4.16

SBIV Pairwise Comparison of Scatter (Range) Between Verbal and Abstract Visual Reasoning Area Scores by Age Subsamples

Value	3 to 6 Yrs %	3 to 6 Yrs Cum %	7 to 11 Yrs %	7 to 11 Yrs Cum %	12 to 23 Yrs %	12 to 23 Yrs Cum %
51 - 55 (-)	0.8	0.8	0.0	0.0	0.0	0.0
46 - 50	0.0	0.8	0.0	0.0	0.0	0.0
41 - 45	0.0	0.8	0.0	0.0	1.1	1.1
36 - 40	0.0	0.8	0.0	0.0	0.0	1.1
31 - 35	0.0	0.8	0.0	0.0	0.0	1.1
26 - 30	1.5	2.3	0.7	0.7	0.0	1.1
21 - 25	4.6	6.9	3.5	4.1	3.3	4.3
16 - 20	3.9	10.8	4.8	8.8	6.5	10.6
11 - 15	9.9	20.8	2.8	11.6	8.5	19.1
6 - 10	4.6	25.4	8.3	19.7	11.7	30.9
1 - 5	10.8	36.2	16.3	36.1	22.4	53.2
0 - 0	1.5	37.7	3.4	39.5	2.1	55.3
1 - 5 (+)	11.5	49.2	16.3	55.8	12.8	68.1
6 - 10	14.7	63.8	15.7	71.4	12.8	80.9
11 - 15	10.8	74.6	10.9	82.3	7.5	88.3
16 - 20	10.0	84.6	7.5	89.8	7.5	95.7
21 - 25	6.9	91.5	3.4	93.2	2.2	97.9
26 - 30	4.7	96.2	2.1	95.2	1.1	98.9
31 - 35	1.5	97.2	2.7	98.0	1.1	100.0
36 - 40	1.6	99.2	1.4	99.3	0.0	100.0
41 - 45	0.0	99.2	0.7	100.0	0.0	100.0
46 - 50	0.0	99.2	0.0	100.0	0.0	100.0
51 - 55	0.0	99.2	0.0	100.0	0.0	100.0
56 - 60	0.8	100.0	0.0	100.0	0.0	100.0
N	130		147		94	
$\bar{X}$	4.48		4.38		-0.12	
SD	16.24		13.28		12.41	
Range (Total)	112.00		70.00		74.00	
Median	06.0		04.0		-01.0	

TABLE 4.17

SBIV Distribution of Minimum Area Score Values by Age Subsamples and Sample

Values	3 to 6 Yrs N	3 to 6 Yrs %	7 to 11 Yrs N	7 to 11 Yrs %	12 to 23 Yrs N	12 to 23 Yrs %	3 to 23 Yrs N	3 to 23 Yrs %
141 - 145	0	0.0	0	0.0	0	0.0	0	0.0
136 - 140	1	0.8	0	0.0	0	0.0	1	0.3
131 - 135	0	0.0	1	0.7	0	0.0	1	0.3
126 - 130	1	0.8	0	0.0	1	1.1	2	0.6
121 - 125	3	2.3	3	2.0	0	2.1	8	2.3
116 - 120	5	3.8	3	2.0	4	4.3	12	3.5
111 - 115	6	4.6	6	4.1	4	4.3	16	4.6
106 - 110	20	15.4	7	4.8	7	7.4	34	9.8
101 - 105	27	20.8	14	9.5	7	7.4	48	13.8
96 - 100	15	11.5	28	19.0	8	8.5	51	14.7
91 - 95	15	11.5	16	10.9	6	6.4	37	10.7
86 - 90	11	8.5	18	12.2	15	16.0	44	12.7
81 - 85	10	7.7	19	12.9	4	4.3	33	9.5
76 - 80	10	7.7	15	10.2	3	3.2	28	8.1
71 - 75	2	1.5	5	3.4	11	11.7	18	5.2
66 - 70	3	2.3	3	2.0	4	4.3	10	2.9
61 - 65	1	0.8	1	0.7	9	9.6	11	3.2
56 - 60	0	0.0	2	1.4	4	4.3	6	1.7
51 - 55	0	0.0	2	1.4	1	1.1	3	0.9
46 - 50	0	0.0	3	2.0	4	4.3	7	2.0
40 - 45	0	0.0	1	0.7	0	0.0	1	0.3
N	130		147		94		371	
$\bar{X}$	97.85		90.73		86.56		92.17	
SD	13.22		15.63		19.73		16.60	
Range	75.00		88.00		80.00		95.00	
Median	100.00		92.00		88.00		94.00	

TABLE 4.18

SBIV Distribution of Maximum Area Score Values by Age Subsamples and Sample

Values	3 to 6 Yrs N	3 to 6 Yrs %	7 to 11 Yrs N	7 to 11 Yrs %	12 to 23 Yrs N	12 to 23 Yrs %	3 to 23 Yrs N	3 to 23 Yrs %
161 - 165	1	0.8	0	0.0	0	0.0	1	0.3
156 - 160	0	0.0	0	0.0	0	0.0	0	0.0
151 - 155	0	0.0	0	0.0	1	1.1	1	0.3
146 - 150	1	0.8	2	1.4	1	1.1	4	1.1
141 - 145	3	2.3	1	0.7	4	4.3	8	2.2
136 - 140	9	6.9	2	1.4	3	3.2	14	3.8
131 - 135	11	8.5	8	5.4	3	3.2	22	5.9
126 - 130	23	17.7	12	8.2	4	4.3	39	10.5
121 - 125	17	12.7	12	8.2	12	12.8	41	11.1
116 - 120	24	17.9	19	12.9	8	8.5	51	13.7
111 - 115	14	10.8	18	12.2	8	8.5	40	10.8
106 - 110	14	10.8	21	14.3	9	9.6	44	11.9
101 - 105	6	4.6	12	8.2	7	7.4	25	6.7
96 - 100	4	3.1	12	8.2	5	5.3	21	5.7
91 - 95	0	0.0	7	4.8	2	2.1	9	2.4
86 - 90	2	1.5	10	6.8	6	6.4	18	4.9
81 - 85	1	0.8	5	3.4	2	2.1	8	2.2
76 - 80	0	0.0	3	2.0	9	9.6	12	3.2
71 - 75	0	0.0	1	0.7	5	5.3	6	1.6
66 - 70	0	0.0	1	0.7	4	4.3	5	1.3
61 - 65	0	0.0	0	0.0	1	1.1	1	0.3
56 - 60	0	0.0	1	0.7	0	0.0	1	0.3
N	130		147		94		371	
$\bar{X}$	120.68		109.74		106.13		112.66	
SD	12.75		16.40		22.18		17.98	
Range	77.00		90.00		90.00		102.00	
Median	120.50		111.00		108.00		115.00	

between Composite Score level and scatter range is not evident (although small statistically significant correlations are present). The finding that is considerable amount of scatter between Area scores is evident in this Clinic sample is important. It is a strong warming to clinicians to be cautious when making interpretations based upon Area score scatter particularly when dealing with specialized populations and when such interpretations may adversely impact examinees. In, brief this section suggests that subtest, Area and Composite Score scatter are now valid areas of research given the introduction of the SBIV and its revised format. It is also suggested that an analysis similar to that of Matarazzo, Daniel, Prifitera and Herman (1988) be conducted on the SBIV standardization sample.

Unrotated Principal Components

Unrotated Principal Components for the 3 to 6 Years Subsample. Table 4.19 displays the unrotated principal component loadings for the youngest subsample and the squared multiple correlations. Typically only two unrotated components are interpreted even though one's theoretical interest may include more. This is because the variance extracted with each successive component is less than the previous component. Therefore, the explanatory power of subsequent components is usually small.

The first component in Table 4.19 is identified by moderately strong positive loadings from all the SBIV subtests. This component, therefore, is identified as a General factor. The second varimax component is biploar with both verbal and nonverbal reasoning themes.

These results can be compared to Table 6.3 of the *Technical Manual'*. A General factor is identified. Salient loadings are obtained on two additional factors from both Verbal and Abstract Reasoning variables. Therefore, both solutions are compatible.

Unrotated Principal Components for the 7 to 11 Years Subsample. Table 4.20 displays the unrotated principal component loadings and squared multiple correlations for the 7 to 11 years subsample. Similar to the youngest Clinic subsample the first component is identified by moderately strong loadgins from all subtests. The second component is

TABLE 4.19

SBIV Unrotated Principal Component Loadings 3 to 6 Years Subsample

	I	II	SMC
Vocabulary	.714	.461	.470
Comprehension	.686	.535	.452
Pattern Analysis	.669	-.465	.365
Copying	.597	-.414	.265
Quantitative	.685	-.371	.364
Bead Memory	.717	-.137	.379
Memory for Sentences	.739	.291	.412
Eigenvalues	3.314	1.128	
Proportion of Total Variance	.473	.161	

TABLE 4.20

SBIV Unrotated Principal Component Loadings 7 to 11 Years Subsample

	I	II	SMC
Vocabulary	.821	-.318	.692
Comprehension	.819	-.253	.659
Absurdities	.747	.054	.528
Pattern Analysis	.801	.297	.618
Copying	.613	.564	.346
Quantitative	.780	.049	.549
Bead Memory	.721	.253	.490
Memory for Sentences	.807	-.318	.654
Memory For Digits	.768	-.236	.591
Memory For Objects	.771	.179	.517
Eigenvalues	5.881	.829	
Proportion of Total Variance	.588	.083	

more moderate than its younger subsample counterpart. However, the loadings are in a similar direction and a bipolar Verbal and Abstract/visual Reasoning component is observed. These results are consistent with but do not necessarily support those in Table 6.5 of the *'Technical Manual'* (p. 57) where three components in addition to the General factor are identified.

Unrotated Principal Components for the 12 to 23 Years Subsample. Table 4.21 provides the unrotated principal components and squared multiple correlations for the oldest subsample. A strong General component is identified. Noteworthy, is the trend for this General component to increase in strength as the subsample ages. A similar pattern is also observed in the standardization data. Unlike the younger two Clinic subsamples the second component is identified by two salient memory variables.

Unrotated Principal Components Summation

There are several findings from the unrotated principal components analysis. First, the presence of a strong General factor is evident. Secondly, the strength of this factor is observed to increase as the subsamples age. The third finding is that both Verbal and Abstract/visual components are observed to emerge in the youngest age groupings. A Short-term memory component is also observed to emerge in the oldest group. These results lend some initial support to the SBIV model. These findings are also supportive of further analyses, particularly those that will separate the influence of the large General factor and highlight potential group factors.

Varimax Rotated Principal Components

A principal components analysis with a varimax rotation of the SBIV subtests is undertaken. As noted in Chapter 3 there are a number of good reasons for using a principal component analysis. Varimax rotations are selected primarily for two reasons. Unlike other orthogonal rotation methods varimax procedures are designed to reduce the opportunity for a General factor to emerge and to maximize the occurrence of group

TABLE 4.21

SBIV Unrotated Principal Component Loadings

for the 12 to 23 Years Subsample

	I	II	SMC
Vocabulary	.880	-.009	.780
Comprehension	.892	-.110	.818
Absurdities	.862	-.207	.780
Pattern Analysis	.829	-.244	.732
Matrices	.897	-.115	.775
Quantitative	.873	-.116	.801
Number Series	.944	-.112	.897
Bead Memory	.825	-.004	.662
Memory For Sentences	.848	.156	.714
Memory For Digits	.809	.304	.663
Memory For Objects	.706	.598	.484
Eigenvalues	8.010	.628	
Proportion of Total Variance	.728	.057	

factors. This is consistent with the SBIV theoretical framework. Secondly, these results can be compared to Sattler's (1988) varimax analysis of the standardization sample data.

Varimax Roations for the 3 to 6 Years Subsample. Table 4.22 displays the two, three and four factor varimax solutions for the youngest sample. Abstract/visual and Verbal reasoning factors semerge in the two component solution. The same components emerge in the three factor solution along with a Visual/spatial reasoning factor. These three components along with a Memory factor emerge when four components are extracted.

These results can be compared to Table 6.3 in the *'Technical Manual'* (p. 55). The Clinic subsample loadings are higher, however, the influence of 'g' has not been extracted. Like the Clinic sample the memory variables in the standardization sample align thesmelves with their respective Reasoning Areas when a Memory factor is not extracted. Also, a Quantitative factor is not found nor expected in either sample. This is because only a single quantitative subtest is available to this age group. Overall, the expected SBIV factor pattern is more readily noticeable in the Clinic sample at this age.

The Clinic subsample results can be more readily and logically compared to the varimax analysis conducted by Sattler (1988) on the standardization sample data. These latter results are supportive of a two component solution at this age according to Sattler. Verbal Comprehension and Nonverbal Reasoning/Visualization components are identified. These results suggest that even under varimax rotations of the standardization data all four cognitive areas do not emerge at this age. Consistent with the Clinic subsample. The Quantitative and Bead Memory subtests align with the Abstract/visual factor. Similarly, Memory for Sentences is aligned with the Verbal reasoning component. Overall, the results of the two varimax analyses are highly similar with the exception of a Short-term memory factor emerging for the Clinic subsample when four components are extracted.

Varimax Rotations for the 7 to 11 Years Subsample. Table 4.23 displays the two, three and four varimax solutions for the middle subsample. Crystalized and Fluid abilities emerge when two components extracted. The three and four component solutions both

TABLE 4.22

SBIV Varimax Rotation for the 3 to 6 Years Subsample

Tests	I	II	h^2	I	II	III	h^2	I	II	III	IV	h^2
Vocabulary	.187	.829	.722	.841	.099	.216	.764	.876	.218	-.018	.147	.187
Comprehension	.115	.862	.757	.846	.209	-.033	.760	.832	.122	.241	-.033	.767
Pattern Analysis	..803	.137	.663	.113	.626	.509	.664	.159	.809	.071	.313	.784
Copying	.716	.123	.528	.170	.188	.910	.894	.122	.246	.141	.925	.950
Quantitative	.749	.215	.608	.152	.804	.208	.713	.163	.799	.324	.050	.773
Bead Memory	.608	.404	.533	.334	.769	.036	.704	.198	.312	.864	.127	.898
Memory for Sentences	.323	.725	.630	.716	.287	.195	.634	.635	.060	.464	.279	.700
Sum of Squared Loadings % of Total Variance	.320	.314		.302	.257	.174		.281	.217	.164	.153	
Cumulative % of Variance	.320	.634		.302	.559	.733		.281	.498	.662	.815	

TABLE 4.23

SBIV Varimax Rotation for the 7 to 11 Years Subsample

Tests	I	II	h^2	I	II	III	h^2	I	II	III	IV	h^2
Vocabulary	.837	.272	.775	.690	.564	.035	.795	.527	.729	.137	.133	.841
Comprehension	.794	.321	.734	.655	.558	.094	.749	.516	.668	.200	.140	.772
Absurdities	.545	.514	.561	.284	.816	.132	.764	.103	.806	.343	.220	.827
Pattern Analysis	.432	.737	.730	.297	.632	.505	.743	.294	.334	.681	.358	.791
Copying	.119	.824	.693	.194	.191	.866	.824	.188	.152	.175	.909	.915
Quantitative	.635	.455	.610	.635	.305	.402	.658	.597	.335	.227	.372	.659
Bead Memory	.399	.652	.583	.225	.667	.379	.639	.270	.222	.869	.122	.892
Memory for Sentences	.826	.264	.752	.803	.319	.183	.780	.777	.335	.293	.083	.809
Memory For Digits	.744	.303	.645	.819	.124	.347	.807	.841	.153	.223	.234	.835
Memory For Objects	.484	.626	.626	.389	.522	.450	.626	.314	.471	.340	.456	.643
Sum of Squared Loadings												
% of Total Variance	.386	.285		.303	.266	.169		.252	.228	.173	.146	
Cumulative % of Variance	.386	.671		.303	.569	.738		.252	.480	.653	.799	

identify Verbal, Abstract and Visual/spatial reasoning abilities. In addition, a complex Short-term memory factor is identified by the four component solution.

These results can be compared to Table 6.4 in the *Technical Manual'* (p. 56). A Quantitative factor is not found in either sample. Secondly, several variables are observed to more closely adhere to the SBIV theoretical model in the Clinic sample (e.g., Copying, Quantitative).

These Clinic subsample results are more readily comparable to Sattler's (1988) varimax of the standardization data. Although Sattler (1988) uses different age groupings in his analysis than the SBIV test authors he generally identifies three components for this age group. These are Verbal Comprehension, Nonverbal reasoning/visualization and Memory. These results are, for the most part, consistent with the Clinic sample. A separate Quantitative component is not identified in either sample. Bead Memory is found to consistently load with Pattern Analysis on an Abstract reasoning component. The Quantitative test, however, is found to load on various components in the Clinic subsample depending on the number of factors extracted. Overall, the two separate varimax results are largely supportive of each other.

Varimax Rotations fo the 12 to 23 Years Subsample. Table 4.24 displays the two, three and four component varimax rotations for the oldest Clinic subsample. These results are relatively more complex than the younger two groups. This is likely due to the stronger 'g' presence. A moderately strong General fcator along with what possibly may be a Short-term memory factor is identified in the two component solution. Crystallized, Fluid and Short-term memory components emerge in the three component solution. The four component solution is more complex, however, it strongly resembles the three component solution as the fourth component is likely a second Memory factor.

These Clinic subsample results can be compared to Table 6.5 in the *Technical Manual'* (p. 57). Here four factors are identified but the Quantitative component is a

TABLE 4.24

SBIV Varimax Rotation for the 12 to 23 Years Subsample

Tests	I	II	h^2	I	II	III	h^2	I	II	III	IV	h^2
Vocabulary	.737	.481	.774	.713	.409	.359	.804	.707	.405	.239	.338	.834
Comprehension	.803	.404	.807	.293	.410	.264	.867	.771	.392	.288	.206	.874
Absurdities	.832	.306	.787	.639	.582	.203	.787	.665	.603	.106	.240	.860
Pattern Analysis	.825	.258	.746	.358	.828	.221	.863	.348	.822	.211	.205	.884
Matrices	.810	.402	.818	.541	.658	.324	.831	.484	.604	.448	.191	.836
Quantitative	.790	.388	.775	.718	.462	.266	.799	.630	.379	.563	.058	.861
Number Series	.847	.430	.903	.652	.609	.328	.903	.596	.557	.450	.197	.907
Bead Memory	.688	.455	.681	.308	.719	.429	.796	.235	.647	.525	.261	.818
Memory for Sentences	.619	.600	.744	.763	.219	.646	.845	.716	.171	.435	.339	.846
Memory For Digits	.505	.702	.747	.439	.384	.641	.751	.320	.262	.769	.354	.887
Memory For Objects	.255	.890	.856	.258	.241	.865	.873	.271	.243	.250	.874	.959
Sum of Squared Loadings % of Total Variance	.521	.264		.349	.286	.194		.307	.250	.185	.128	
Cumulative % of Variance	.521	.785		.349	.635	.829		.307	.557	.742	.870	

singlet. This is conistent with the Clinic varimax results as a Quantitative factor is also not present. The remaining factors are consistent across both groups.

The Clinic subsample varimax results are more readily comparable to Sattler's (1988) varimax analysis. Abstract reasoning and Short-term memory components are identified in both analyses. Similarly a separate Quantitative component does not emerge in either group. The quantitative variables studied have salient loadings on both Verbal and Abstract reasoning factors. However, loadings are highest on the former lending to a possible Crystallized factor interpretation. Overall, the two varimax analyses are largely supportive of each other.

Varimax Rotated Principal Components Summation

Varimax procedures are designed to maximize the opportunity for separate or multiple factors. This is a favored procedure given the SBIV theoretical structure. This procedure is effective in identifying three components across the three subsamples. These are Verbal reasoning, Abstract or Visual-spatial reasoning and Short-term memory. A Quantitative factor did not emerge. These results are generally consistent with the *Technical Manual'* factors and a subsequent varimax analysis conducted by Sattler (1988).

Due to the presence of a proposed theoretical structure a more logical approach than Sattler's exploratory varimax techniques is confirmatory analysis. Sattler also appears to rotate four components and then he only interprets three. Therefore, one may question Sattler's choice of methods and his unwillingness to confirm or deny the existence of the SBIV theoretical structure.

Procrustes Rotations

Orthogonal Procrustes Rotations for the 3 to 6 Years Subsample. Tables 4.25 to 4.28 display the two, three and four orthogonal target and Procrustean rotations for the youngest subsample. Relative to the oblique Procrustes rotations the orthogonal is more rigorous. That is, orthogonal rotations are less susceptible to chance findings. As a result

TABLE 4.25

SBIV Two Factor Target and Orthogonal Procrustes Solutions

for the 3 to 6 Years Subsample

Variable	Factor I		Factor II		h^2
	Target	Rotated	Target	Rotated	
Vocabulary	1.000	.847	.000	.063	.721
Comprehension	1.000	.870	.000	-.014	.757
Pattern Analysis	.000	.254	1.000	.774	.664
Copying	.000	.227	1.000	.690	.528
Quantitative	1.000	.323	.000	.709	.607
Bead Memory	.000	.490	1.000	.541	.533
Memory for Sentences	1.000	.765	.000	.212	.630
Sum of Squares		2.520		1.920	
Proportion of Total Variance		.360		.274	
Cumulative Variance		.360		.634	
Coefficients of Congruence		.88		.84	

they are the rotation of choice. A more complete discussion of these techniques is provided in Chapter 3.

The two factor orthogonal Procrustes solution is displayed in Table 4.25. The first factor is identified as a Crystallized ability. Large salient loadings are obtained from both verbal and quantitative variables. The congruence coefficient is .88. This factor is able to account for 36 percent of the matrix variance.

The second factor is identified as a Fluid or Abstract/visual component. Salient loadings are obtained from all variables with the exception of the three language based tests. Noteworthy, is that both Bead Memory and Quantitative test are salients on both factors. However, relatively stronger loadings are observed on the second factor. The congruence coefficient for this second factor is .84 and it is able to account for 27 percent of the matrix variance. Overall, the two factor solution is able to account for 63 percent of the matrix variance.

The three component orthogonal Procrustes solution for the youngest subsample is displayed in Table 4.26. Verbal, Abstract/visual and Visual/Spatial abilities are identified. They are respectively able to account for 32, 21 and 20 percent of the matrix variance. Overall, this solution is able to account for 73 percent of the matrix variance. Congruence coefficients are in the range of .59 to .89.

The four factor orthogonal Procrustes solution for this age group is displayed in Table 4.27. The first factor is identified as a Verbal ability. It is able to explain 24 percent of the matrix variance. The congruence coefficient is .90.

The second factor of this solution is likely an Abstract/Visual ability. It is able to explain 22 percent of the matrix variance. The congruence coefficient is .89.

The third factor is defined by three salients. These are Bead Memory, Pattern Analysis, and Quantitative. The relative strength of these loadings in comparison to the second component, suggest higher level abstract reasoning processes are involved. Relative to other salients, the highest loading is observed on the Quantitative test. This

TABLE 4.26

SBIV Three Factor Target and Orthogonal Procrustes Solutions for the 3 to 6 Years Subsample

Variable	Factor I		Factor II		Factor III		h^2
	Target	Rotated	Target	Rotated	Target	Rotated	
Vocabulary	1.00	.861	0.00	.128	0.00	.079	.764
Comprehension	1.00	.830	0.00	-.077	0.00	.256	.760
Pattern Analysis	0.00	.171	1.00	.650	0.00	.460	.664
Copying	0.00	.290	1.00	.897	0.00	-.069	.894
Quantitative	1.00	.166	0.00	.411	0.00	.719	.713
Bead Memory	0.00	.323	0.00	.214	1.00	.744	.704
Memory For Sentences	0.00	.731	0.00	.177	1.00	.259	.634
Sum of Squares		2.210		1.495		1.426	
Proportion of Total Variance		.316		.214		.204	
Cumulative Variance		.316		.530		.734	
Congruence Coefficients		.72		.89		.59	

TABLE 4.27

SBIV Four Factor Target and Orthogonal Procrustes Solutions for the 3 to 6 Years Subsample

Variable	Factor I		Factor II		Factor III		Factor IV		h^2
	Target	Rotated	Target	Rotated	Target	Rotated	Target	Rotated	
Vocabulary	1.00	.866	0.00	.233	0.00	.059	0.00	.172	.837
Comprehension	1.00	.778	0.00	.002	0.00	.099	0.00	.389	.767
Pattern Analysis	0.00	.186	1.00	.646	0.00	.569	0.00	.091	.784
Copying	0.00	.049	1.00	.910	0.00	.099	0.00	.292	.950
Quantitative	0.00	.158	0.00	.383	1.00	.718	0.00	.295	.773
Bead Memory	0.00	.044	0.00	.162	0.00	.339	1.00	.869	.898
Memory For Sentences	0.00	.517	0.00	.221	0.00	-.044	1.00	.618	.700
Sum of Squares		1.686		1.522		1.003		1.499	
Proportion of									
Total Variance		.241		.217		.143		.214	
Cumulative Variance		.241		.458		.601		.815	
Congruence Coefficients		.90		.89		.72		.86	

factor is able to explain 14 percent of the variance in the matrix. The congruence coefficient is .72.

The fourth factor is largely memory based. It is able to explain 21 percent of the matrix variance. The congruence coefficient is .86. Overall, the four component orthogonal Procrustes solution is able to account for 82 percent of the matrix variance. Therefore, the four component solution for this age group is best able to approximate the proposed SBIV model. Relative to the other solutions, it is also able to account for the largest proportion of variance in the matrix.

These results can be compared to Table 6.3 in the *Technical Manual'*. Only two components, in addition to the General factor, are extracted by the test authors for this age group. Comparing these results to the two component Procrustes solution is informative The first factor in both samples is clearly Verbal. Clinic sample loadings are higher. Additional salient loadings are obtained on this factor in the Clinic subsample.

It is the second Clinic subsample factor that fares relatively better than its standardization sample counterpart. Four of the seven variables are salients on this second factor in the Clinic data. The abstract reasoning qualities in this second factor are also more readily apparent in the Clinic data.

Relative to both standardization sample and the Clinic Orthogonal Procrustes subsample data the four component Procrustes solution is the best approximation of the SBIV model. Verbal, Abstract and Memory factors are all evident. Some support is also found for a Quantitative or higher level Abstract reasoning component. This solution is also able to account for the largest proportion of the total matrix variance.

Orthogonal Procrustes Rotations for the 7 to 11 Years Subsample. Tables 4.28 to 4.30 provide the two, three and four orthogonal target and Procrustes rotations for the middle subsample. Similar to the unrotated principal components results, the influence of a strong General factor is more apparent as the subsamples age. That is, a greater number of variables are observed to load on more than one factor in each of the solutions.

TABLE 4.28

SBIV Two Factor Target and Orthogonal Procrustes Solutions

for the 7 to 11 Years Subsample

Variable	Factor I		Factor II		h^2
	Target	Rotated	Target	Rotated	
Vocabulary	1.000	.859	.000	.199	.774
Comprehension	1.000	.821	.000	.244	.734
Absurdities	1.000	.592	.000	.459	.561
Pattern Analysis	.000	.501	1.000	.692	.730
Matrices	.000	.197	1.000	.809	.693
Quantitative	1.000	.675	.000	.392	.610
Bead Memory	.000	.459	1.000	.610	.583
Memory for Sentences	1.000	.848	.000	.184	.753
Memory for Digits	1.000	.769	.000	.230	.644
Memory for Objects	.000	.542	1.000	.576	.626
Sum of Squares		4.323		2.385	
Proportion of Total Variance		.432		.239	
Cumulative Variance		.432		.671	
Coefficients of Congruence		.90		.87	

TABLE 4.29

SBIV Three Factor Target and Orthogonal Procrustes Solutions for the 7 to 11 Years Subsample

Variable	Factor I		Factor II		Factor III		h^2
	Target	Rotated	Target	Rotated	Target	Rotated	
Vocabulary	1.00	.730	0.00	-.018	0.00	.512	.795
Comprehension	1.00	.698	0.00	.042	0.00	.510	.749
Absurdities	1.00	.800	0.00	.314	0.00	.160	.764
Pattern Analysis	0.00	.540	1.00	.554	0.00	.380	.743
Copying	0.00	.013	1.00	.730	0.00	.539	.824
Quantitative	1.00	.381	0.00	.214	0.00	.683	.658
Bead Memory	0.00	.581	0.00	.487	1.00	.252	.639
Memory for Sentences	0.00	.506	0.00	-.028	1.00	.723	.780
Memory for Digits	0.00	.293	0.00	.030	1.00	.849	.807
Memory for Objects	0.00	.485	0.00	.430	1.00	.454	.626
Sum of Squares		3.012		1.411		2.962	
Proportion of Total Variance		.301		.141		.296	
Cumulative Variance		.301		.442		.738	
Congruence Coefficients		.75		.76		.66	

TABLE 4.30

SBIV Four Factor Target and Orthogonal Procrustes Solutions for the 7 to 11 Years Subsample

Variable	Factor I Target	Factor I Rotated	Factor II Target	Factor II Rotated	Factor III Target	Factor III Rotated	Factor IV Target	Factor IV Rotated	[illegible]
Vocabulary	1.00	.753	0.00	.043	0.00	.248	0.00	.459	.841
Comprehension	1.00	.688	0.00	.089	0.00	.215	0.00	.494	.772
Absurdities	1.00	.794	0.00	.326	0.00	-.106	0.00	.280	.827
Pattern Analysis	0.00	.320	1.00	.495	0.00	-.122	0.00	.654	.781
Copying	0.00	.139	1.00	.834	0.00	.392	0.00	.216	.915
Quantitative	0.00	.356	0.00	.270	1.00	.367	0.00	.570	.659
Bead Memory	0.00	.204	0.00	.350	0.00	-.356	1.00	.775	.892
Memory For Sentences	0.00	.370	0.00	-.015	0.00	.309	1.00	.759	.809
Memory for Digits	0.00	.192	0.00	.071	0.00	.470	1.00	.757	.835
Memory for Objects	0.00	.468	0.00	.465	0.00	.147	1.00	.431	.643
Sum of Squares		2.354		1.474		.886		3.270	
Proportion of Total Variance		.235		.147		.089		.327	
Cumulative Variance		.235		.382		.471		.798	
Congruence Coefficients		.84		.77		.39		.75	

The two factor orthogonal Procrustes solution for the 7 to 11 years subsample is displayed in Table 4.28. The first factor is identified as a Crystallized ability as relatively stronger loadings are observed on verbal and quantitative variables. However, the strong influence of a General factor is also present given the number of salient loadings. The congruence coefficient for this first component is .90. This factor is able to account for 43 percent of the total matrix variance.

The second factor is identified as a Fluid ability. Relatively higher loadings are observed from tests that possess abstract reasoning qualities. Loadings are also observed to cross over the cognitive reasoning areas of ability as outlined by Thorndike, Hagen, and Sattler. Therefore, this factor is perceived to possess a greater scope of influence. The congruence coefficient for this second component is .87. It is able to account for 24 percent of the matrix variance. This solution is able to explain 67 percent of the total variance.

The three factor orthogonal Procrustes solution for the 7 to 11 years subsample is displayed in Table 4.29. Eight of ten variables are salients on the first factor. This is also true for the third component. Relatively higher loadings are observed from language relative variables on the first factor. Whereas, memory variables are somewhat stronger on the third component. The second component is defined by variables with abstract/visual themes. All three components ranging from .66 to .75. A total of 74 percent of the matrix variance is explained by this solution.

Table 4.30 displays the four factor orthogonal Procrustes solution for the middle Clinic subsample. Similar to the middle subsample the majority of variables have salient loadings on the first and last components. Relatively higher loadings are observed on verbal variables on the first factor. Whereas memory variables are relatively larger on the fourth component. The second and third components are respectively identified as Abstract/Visual and Quantitative factors. The latter is least well defined. Congruence coefficients range from .39 (Quantitative) to .84 (Verbal). This four component solution is able to explain approximately 80 percent of the total variance in the matrix.

These results can be compared to Table 6.4 of the *Technical Manual*'. Three factors, in addition to a strong General factor, are identified in the standardization data. These are Verbal, Memory, and Abstract/Visual. Importantly, the latter is a singlet. As noted earlier some variables fail to load on their specified factor (e.g., Absurdities) and others load on factors outside their cognitive area (e.g., Bead Memory).

Both two and four component orthogonal solutions are perceived to be more clear approximations of the proposed SBIV theoretical models. The two component solution is generally defined along the Crystallized-Fluid model advocated by the test authors. The four cognitive ability areas are more readily evident in the four component orthogonal Procrustes solution. However, like the standardization data a Quantitative factor is least well defined. In addition, a number of variables in the four component Clinic solution have salient loadings on more than one component.

Orthogonal Procrustes Rotations for the 12 to 23 Years Subsample. Tables 4.31 to 4.33 display the two, three and four orthogonal target and Procrustes rotations for the oldest Clinic subsample. The influence of a strong General factor is more apparent here, than in the two younger groups. The first factor of the two component solution is largely general and accounts for approximately 58 percent of the total variance. Nine of the second factor variables are also salient. The congruence coefficients for these two factors are .81 and .58 respectively. This solution is able to explain approximately 79 percent of the matrix variance.

The first and third components of the three component solution are relatively strong General factors. The second factor is identified as an Abstract/Visual component. Congruence coefficients are moderately low ranging from .68 to .79. This solution is able to account for 83 percent of the total variance in the matrix.

TABLE 4.31

Orthogonal Procrustes Rotations for the 12 to 23 Years Subsample

Variable	Factor I		Factor II		h^2
	Target	Rotated	Target	Rotated	
Vocabulary	1.000	.782	.000	.403	774
Comprehension	1.000	.840	.000	.319	.808
Absurdities	1.000	.859	.000	.219	.786
Pattern Analysis	.000	.847	1.000	.172	.747
Matrices	.000	.847	1.000	.317	.817
Quantitative	1.000	.825	.000	.306	.775
Number Series	1.000	.887	.000	.342	.904
Bead Memory	.000	.731	1.000	.382	.680
Memory for Sentences	1.000	.678	.000	.534	.745
Memory for Digits	1.000	.574	.000	.647	.748
Memory for Objects	.000	.344	1.000	.859	.856
Sum of Squares		6.339		2.241	
Proportion of Total Variance		.582		.204	
Cumulative Variance		.582		.786	
Coefficients of Congruence		.81		.58	

TABLE 4.32

SBIV Three Factor Target and Orthogonal Procrustes Solutions for the 12 to 23 Years Subsample

Variable	Factor I Target	Factor I Rotated	Factor II Target	Factor II Rotated	Factor III Target	Factor III Rotated	h^2
Vocabulary	1.00	.762	0.00	.153	0.00	.447	.804
Comprehension	1.00	.849	0.00	.140	0.00	.357	.867
Absurdities	1.00	.754	0.00	.354	0.00	.307	.787
Pattern Analysis	0.00	.552	1.00	.667	0.00	.335	.863
Matrices	0.00	.671	1.00	.443	0.00	.430	.831
Quantitative	1.00	.790	0.00	.211	0.00	.360	.799
Number Series	1.00	.762	0.00	.364	0.00	.435	.903
Bead Memory	0.00	.455	0.00	.559	1.00	.526	.796
Memory For Sentences	0.00	.748	0.00	-.052	1.00	.532	.845
Memory for Digits	0.00	.467	0.00	.182	1.00	.706	.751
Memory for Objects	0.00	.235	0.00	.077	1.00	.900	.973
Sum of Squares		4.870		1.341		2.908	
Proportion of Total Variance		.443		.122		.264	
Cumulative Variance		.443		.565		.829	
Coefficients of Congruence		.79		.68		.78	

TABLE 4.33

SBIV Four Factor Target and Orthogonal Procrustes Solutions for the 12 to 23 Years Subsample

Variable	Factor I		Factor II		Factor III		Factor IV		h^2
	Target	Rotated	Target	Rotated	Target	Rotated	Target	Rotated	
Vocabulary	1.00	.683	0.00	.223	0.00	.281	0.00	.489	.834
Comprehension	1.00	.691	0.00	.205	0.00	.420	0.00	.422	.974
Absurdities	1.00	.710	0.00	.435	0.00	.237	0.00	.338	.860
Pattern Analysis	0.00	.440	1.00	.716	0.00	.212	0.00	.336	.884
Matrices	0.00	.420	1.00	.476	0.00	.427	0.00	.501	.836
Quantitative	0.00	.429	0.00	.234	1.00	.635	0.00	.468	[illegible]
Number Series	0.00	.509	0.00	.406	1.00	.470	0.00	.512	.907
Bead Memory	0.00	.196	0.00	.569	0.00	.336	1.00	.586	.808
Memory For Sentences	0.00	.567	0.00	-.004	0.00	.406	1.00	.600	.846
Memory for Digits	0.00	.115	0.00	.172	0.00	.462	1.00	.795	.887
Memory for Objects	0.00	.371	0.00	.130	0.00	-.224	1.00	.868	.959
Sum of Squares		2.772		1.606		1.704		3.486	
Proportion of Total Variance		.252		.146		.155		.317	
Cumulative Variance		.252		.398		.553		.870	
Coefficients of Congruence		.72		.66		.60		.76	

The four component orthogonal Procrustes solution for the oldest subsample is outlined in Table 4.33. The effects of a strong General component are, again, evident. Multiple salient loadings are observed on the first, third and fourth components. Verbal loadings are relatively higher on the first factor. Quantitative variables are stronger on the third component. Memory variables are relatively stronger on the fourth component. The second component is identified as an Abstract/Visual factor. Overall, this solution is able to explain 87 percent of the total variance.

The above results can be compared to Table 6.5 in the *Technical Manual'*. General, Verbal, Memory, Abstract/Visual and Quantitative factors are identified. Importantly, the latter factor is a singlet. Despite the fact that the two remaining quantitative variables are nonsalients, it is evident that these results more strongly support the SBIV theoretical model. Therefore, the four cognitive areas of ability, as delineated by Thorndike, Hagen and Sattler (1986(a)(b)) are more readily observed in the standardization sample results for this age group.

Orthogonal Procrustes Summation

In summation the Procrustes results are complex and provide only mixed support for the proposed SBIV theoretical structure. Moderate support is found for Crystallized and Fluid abilities in the two younger subsample two component solutions. Congruence coefficients are in .80 to .90 range. Moderate support for the presence of a Crystallized ability factor is also found in the three component solutions. Congruence coefficients are in the low to high seventies and increase with age.

The Fluid ability factor for the three component solution is strongest for the youngest group (.89) and weakest for the older group (.68). The Short-term memory factor for the three component solution is the most tenuous. Congruence coefficients range from .59 to .78 and increase with age.

Individual test loadings results are also complex for both the two and three component solutions. The majority of tests have salient loadings on more than one factor and the patterns of complex loadings change between age groups. Some tests appear to lack specificity at certain age groups as they have similar salient loadings across all three components (e.g., Memory for Objects for the middle subsample and Bead Memory for the oldest subsample). Memory for Sentences has higher loadings on the Crystallized ability as opposed to Short-term memory factor for two of the three subsamples. Overall, the memory tests appear to be the least differentiated as they typically have salient loadings on more than one factor.

The four factor Procrustes solutions are also complex and many of the difficulties noted in the three component solutions are relevant. They most closely adhere to the proposed SBIV model. However, there are some interpretative differentiations at various ages. Several variables are also observed to load on more than one factor.

Strongest support is found for a Verbal factor. Factor loadings and congruence coefficients, on average, are highest for this variable. The congruence coefficient is inversely related to age for this first factor.

The Short-term memory factor is better defined in the four component solutions. Coefficients are typically higher relative to the three component solutions. However, the Memory tests are consistently complex, loading on more than one factor. Bead Memory is particularly complex for the oldest subsample. Memory for Objects for the middle subsample is complex showing approximately equivalent loadings on three factors. The congruence coefficients range from .75 to .86.

The Abstract/visual factor is moderately strong and like the Fluid Factor in the three component solution it decreases with age. Pattern Analysis typically has salient loadings on more than one factor. For the middle subsample its highest loading is on the Short-term memory factor. More consistent loadings are found on the other tests designated to describe this factor. The congruence coefficient for the Abstract/visual factor ranges from .56 to .89.

The Quantitative factor is the least well defined of the four components. A possible explanation for this finding is that for the youngest group only one possible test is available to describe this factor. Although other tests are available for the middle subsample, only one is used with the frequency required to be included in this analysis. Two of the three available Quantitative tests are used in the oldest subsample analysis. Overall, the Quantitative test loadings are low to moderately low. For the Middle subsample a higher loading is found on the Short-term memory factor. A similar problem is found for the oldest group where the Number Series test has higher loadings on both the Verbal and Short-term memory factors. However, the Number Series test attained salient loadings on all four factors for this age group possibly suggesting a lack of factor specificity. The congruence coefficients are in the range of .75 to .86.

Systematic Permutations

Systematic permutations are used as a subsidiary analysis to the orthogonal Procrustes rotations. They are employed to further clarify relationships between variables. All possible orthogonal hypotheses of a similar form are utilized as a reference set. The determination of statistical significance (or the probability value) is based on 5,040 systematic permutations of the rows of the target matrix. Therefore, all possible permutations are examined. The null hypothesis is that all permutations of the hypothesis matrix are equally likely for this data. The P-value (significance or probability) value is the proportion of data permutations that have Trace (E'E) values less than or equal to the results obtained under the hypothesized structures. E is the hypothesis matrix minus the rotated matrix. In otherwords E is the error matrix. Trace (E'E) is simply the sum of squares of E.

The trace significance values for each subsample for the two (i.e., Crystallized, Fluid), three (i.e., Crystallized, Fluid and Short-term memory) and four (i.e., Verbal Reasoning, Abstract/visual Reasoning, Quantitative Reasoning and Short-term memory) factor solutions are displayed in Table 4.34. At the .05 level of significance the two factor

TABLE 4.34

Systematic Permutation Statistical Significance Levels by Subsample

Subsample	II Factor Model	III Factor Model	IV Factor Model
3 to 6 years	.086	.357	.021
7 to 11 years	.004	.172	.049
12 to 23 years	.608	.003	.015

model is significantly better at explaining the obtained results than random models at the 7 to 11 years age level. The youngest subsample, at .086 is, however, close to approximating significance. The two factor model fits least well for the oldest subsample.

The three factor model, relative to systematic data arrangements, is a statistically significant better explanation of the obtained results for the oldest subsample. This is not true for the two younger subsamples. Both two and four factor models are better able to explain the findings than the three factor models for these latter groups.

For all three subsamples the four factor model is a better explanation of the Clinic data than systematically random structures. All three subsamples are able to meet the .05 level of significance. This solution is a relatively better fit for the oldest subsample.

In sum, the systematic permutation results are not uniform across the three subsamples. The two factor model is best able to account for the 7 to 11 years subsample data. The best model for the 12 to 23 years subsample, according to this analysis, is the 3 factor model. The 4 factor model is a relatively better explanation for the 3 to 6 years subsample data.

Overall, the four factor model is superior to both two and three factor interpretation at this level. That is, it is best able to account for the obtained data across all three age groups.

Ancillary Orthogonal Procrustes Analysis

Three additional criteria have been proposed by Helmes (1989) to assist in evaluating the internal structure of psychometric measures. These criteria also use targeted rotations and the hypothesized theoretical structures. They are based on the number of items (tests) having their highest loading on their keyed (area) scale, the mean loading of keyed items (tests) and the number of items (tests) having their highest loading on non-keyed (areas) scales. These criteria are said to avoid the difficulties encountered by comparing various rotational methods on the same data and debates over an acceptable minimum factor loading. These criteria also allow for the examination of each hypothesized factor in the various solutions.

The proportion of items (tests) having their highest loading on their hypothesized factor has high intuitive value. Instruments that have good internal structure should have tests that load most highly on their designated factor. Ideally, this proportion (P_k) should be 1.0. Proportions that are below .5 are generally viewed as being unacceptable. The intermediate range is defined as being between .5 and .8.

The mean loading of keyed items (M_k) is a measure of scale or area content saturation. Helmes (1989) uses Comrey's (1978) qualitative labels to evaluate the size of loadings. These are: excellent (.71), very good (.63), good (.55), fair (.45) and poor (.32). An M_k over .7 is desirable.

The number of items having their highest loading on non-keyed items (Nnk) is a measure of discriminate validity. It is based on the notion that variables should only load on areas which they are hypothesized or keyed. An Nnk of .0 is desirable.

Table 4.35 outlines the results from each of the three criteria by subsample and number of factors.

An analysis of the two factor solutions suggests that this model is more appropriate for the 7 to 11 years subsample. All three criteria are basically met. This result is consistent with the randomization test data. The primary problem for the younger group is the higher loading of the Quantitative variable on the Fluid as opposed to Crystallized factor.

The two factor model is least acceptable for the oldest subsample. The primary problem for the first factor of this solution is the number of non-keyed tests that have their highest loading on this factor. Pattern Analysis, Matrices and Bead Memory all have higher loadings on this hypothesized Crystallized ability. Both Bead Memory and Matrices are also salients on their keyed factor.

The primary difficulty for the second factor is the number tests having their highest loading on their designated factor. Of the four tests designated to load on this factor only

TABLE 4.35

The Proportion of Items Loading Higher on their Keyed Scale than on any other, Mean Keyed Loading, and the Proportion of Intruding Non-Keyed Items for SBIV Factor Models by Subsample

3 to 6	I	II	I	II	III	I	II	III	IV
P_k	.750	1.000	.667	1.000	.500	1.000	1.000	1.000	1.000
M_k	.701	.669	.619	.773	.502	.822	.718	.778	.744
N_k	.000	.250	.250	.200	.200	.000	.000	.000	.000
7 to 11	I	II	I	II	III	I	II	III	IV
P_k	1.000	1.000	.750	1.000	.500	1.000	.000	.500	.750
M_k	.761	.672	.652	.642	.569	.745	.367	.664	.681
N_k	.000	.000	.333	.125	.167	.143	.000	.125	.333
12 to 23	I	II	I	II	III	I	II	III	IV
P_k	.857	.250	1.000	.500	.500	1.000	.500	.500	1.000
M_k	.778	.443	.783	.555	.666	.694	.553	.596	.712
N_k	.750	.143	.333	.111	.000	.125	.000	.000	.286

Memory for Objects is able to meet this criteria. Pattern Analysis, Matrices and Bead Memory have higher loadings on the first factor.

The three factor model is most appropriate for the oldest subsample according to the randomization test data. Differences between subsamples are more subtle using the three criteria. All factors from the three subsamples meet the number of tests having their highest loading on their hypothesized area.

Some tests, however, have highest loadings on non-keyed factors. For the two younger subsamples the Quantitative test loads highest on the Short-term memory factor. This suggests that at the younger ages quantitative skills may have a large memory component.

The second factor is less affected by tests having higher loadings on their non-keyed factors. It is only in the oldest group where Matrices is observed to load more strongly on the Crystallized factor. All other tests are observed to load relatively more highly on their designated factors.

It is the third factor where this criteria is least well met. For both the youngest and oldest subsamples Memory for Sentences has its highest loading on the Crystallized ability. Bead Memory is the least consistent in meeting this criteria. For the middle group its highest loading is on the Crystallized factor. For the oldest group its highest loading is on Fluid ability. Like Bead Memory, Memory for Objects has its highest loading on the Crystallized component.

The only component in the four factor solutions that is seriously unable to meet the P_k and M_k criteria is found in the middle group. This is the Quantiative factor. The Quantitative component is, however, defined by a single variable at this age level. Therefore, its deviation is less worrisome.

Several more variables, however, have their highest loading on non-keyed items. This problem is not relevant to the youngest subsample. For the middle subsample both Quantitative and Pattern Analysis have their highest loadings on the Short-term memory factor. Memory for Objects has its highest loading on the Verbal factor. For the oldest subsample Matrices attains its highest loading on the Short-term memory component. Number Series has its highest loading on the Verbal factor.

An analysis of the variance explained by each variable is often a useful exercise. Such an analysis is taken a step further through the separation of content and nuisance variance. Content variance is defined as the amount of variance accounted for by a designated factor. Nuisance variance is the amount of variance explained by nondesignated factors. Nuisance variance is said to predominate when it accounts for a greater proportion of the total variance than content variance.

The communalities of each variable by both age group and number of factors extracted are displayed in Tables 4.36 to 4.38. The total variance accounted for by each variable is divided into both content and nuisance variance. It is evident that in the vast majority of cases content variance exceeds nuisance variance.

The Nuisance variance predominates for the Quantitative variables in at least two solutions at each age level. Nuisance variance also predominates in the three and four factor solutions at the 7 to 11 year level for Pattern Analysis and Memory for Objects. At the 12 to 23 year level, nuisance variance is relatively greater in all three solutions for Matrices and Bead Memory. Nuisance variance is also greater than content variance for Memory for Sentences, at this age level, in two of the three solutions. These results are consistent with the three criteria analysis.

Ancillary Orthogonal Procrustes Analysis Summation

The above analyses are generally supportive of a four factor interpretation of the Clinic data that largely coincides with the four cognitive areas of ability. Specific submodels are observed to be a somewhat better for both the middle and older subsamples. However, the four factor model is able to garnish support across all three age groups. Helmes' (1989) criteria is largely supportive of the four component model. It warns against factor interpretations that are based on single variables as their relationships both

TABLE 4.36

SBIV Communalities by Number of Factors Extracted for the 3 to 6 Years Subsample

Variables	II Factor Model			III Factor Model			IV Factor Model		
	Total Variance	Content Variance	Nuisance Variance	Total Variance	Content Variance	Nuisance Variance	Total Variance	Content Variance	Nuisance Variance
Vocabulary	.722	.718	.004	.764	.741	.023	.838	.750	.087
Comprehension	.757	.757	.000	.760	.689	.071	.767	.605	.161
Pattern Analysis	.663	.599	.065	.664	.422	.241	.784	.418	.366
Copying	.528	.477	.052	.893	.804	.089	.951	.828	.122
Quantitative	.608	.105	.503*	.713	.028	.686*	.773	.515	.258
Bead Memory	.533	.293	.240	.704	.554	.150	.898	.755	.143
Memory for Sentences	.630	.585	.045	.634	.067	.566*	.700	.383	.318

* Signifies Nuisance Variance Predominates

TABLE 4.37

SBIV Communalities by Number of Factors Extracted for the 7 to 11 Years Subsample

Variables	II Factor Model			III Factor Model			IV Factor Model		
	Total Variance	Content Variance	Nuisance Variance	Total Variance	Content Variance	Nuisance Variance	Total Variance	Content Variance	Nuisance Variance
Vocabulary	.775	.739	.036	.795	.533	.262	.841	.567	.274
Comprehension	.734	.675	.059	.749	.487	.262	.772	.474	.298
Absurdities	.561	.350	.211	.764	.639	.124	.827	.630	.196
Pattern Analysis	.730	.479	.251	.743	.307	.436*	.791	.245	.545*
Copying	.693	.655	.039	.824	.534	.291	.915	.695	.220
Quantitative	.610	.456	.154	.658	.146	.512*	.660	.135	.525*
Bead Memory	.583	.372	.211	.639	.064	.575*	.892	.601	.291
Memory for Sentences	.752	.719	.034	.780	.522	.257	.809	.577	.233
Memory for Digits	.645	.592	.053	.807	.721	.087	.835	.572	.263
Memory for Objects	.626	.332	.294	.626	.206	.420*	.643	.186	.457*

* Signifies Nuisance Variances Predominates

TABLE 4.38

SBIV Communalities by Number of Factors Extracted for the 12 to 23 Years Subsample

Variables	II Factor Model			III Factor Model			IV Factor Model		
	Total Variance	Content Variance	Nuisance Variance	Total Variance	Content Variance	Nuisance Variance	Total Variance	Content Variance	Nuisance Variance
Vocabulary	.774	.612	.162	.804	.581	.223	.834	.466	.368
Comprehension	.807	.705	.102	.867	.720	.147	.874	.478	.397
Absurdities	.787	.738	.048	.787	.568	.219	.860	.503	.357
Pattern Analysis	.746	.030	.717*	.863	.445	.418	.885	.512	.372
Matrices	.818	.101	.717*	.831	.196	.634*	.836	.227	.610*
Quantitative	.775	.681	.093	.799	.624	.175	.861	.403	.457*
Number Series	.903	.786	.117	.903	.581	.322	.907	.221	.686*
Bead Memory	.681	.146	.535*	.796	.277	.519*	.818	.343	.475*
Memory for Sentences	.744	.460	.285	.845	.283	.562*	.846	.361	.486*
Memory for Digits	.748	.329	.418*	.751	.499	.252	.887	.631	.256
Memory for Objects	.856	.738	.119	.873	.812	.061	.959	.754	.205

* Signifies Nuisance Variance Predominates

within and between factors may change. Similar to earlier analyses the influence of a General factor is relatively stronger in older subjects.

Multiple-Group Analysis

Oblique Multiple-Group Analysis

As outlined in Chapter 3 Multiple group factor analysis is considered an excellent approach for determining the existence of single or multiple factor solutions (Nunnally, 1967; For detailed discussions of this procedure see Gorsuch, 1983; Harman, 196[illegible] McDonald, 1985). Multiple-group factor analysis is well accepted procedure for [illegible] subscale items within tests. It is an especially valuable [illegible] for investigating theoretical structures that are purported to exist within a test battery. Oblique multiple group analysis is particularly valuable in this analysis as it closely parallels the methods employed by Thorndike, Hagen and Sattler (1986b) on the standardization sample.

Oblique Multiple-Group Analysis for the 3 to 6 Years Subsample. Tables 4.39 to 4.41 display the two, three and four oblique multiple-group solutions for the 3 to 6 years subsample. Crystallized and Fluid abilities are readily identified in the two component solution. Congruence coefficients of .88 and .83 are obtained.

The three component solution is in line with the two component results. Crystalized, Fluid and Memory abilities are defined. Congruence coefficients are somewhat more moderate ranging from .75 to .80.

The four component solution is observed to closely follow the SBIV hypothetical model. The first factor is clearly Verbal. The second component is defined by a single loading that reaches unity. This component is tentatively identified as a Quantitative factor. The absence of additional quantitative variables for this age group is a major impediment to the definition of this component. The third and fourth components are delineated as Abstract/visual and Short-term memory components. Congruence coefficients are moderate and range from .72 to .82.

TABLE 4.39

Oblique Multiple-Group Factor (Pattern) Analysis for the 3 to 6 Years Subsample

Variables	I	II	I	II	III	I	II	III	IV
Vocabulary	.901	-.143	.897	-.056	-.058	.906	.020	.057	-.057
Comprehension	.879	-.152	.860	-.177	.057	.892	-.020	-.057	.057
Pattern Analysis	-.034	.823	.070	.834	-.014	.013	.100	.801	-.014
Copying	-.070	.834	-.070	.877	.014	-.013	-.100	.899	.014
Quantitative	.469	.297	.574	.233	.001	.000	1.00	.000	.000
Bead Memory	.143	.674	-.094	.067	.880	-.138	.071	.033	.880
Memory for Sentences	.792	-.003	.094	-.067	.822	.138	-.071	-.033	.823
Sum of Squares	2.465	1.958	1.902	1.545	1.458	1.654	1.031	1.459	1.458
Coefficients of Congruence	.88	.83	.80	.80	.75	.82	.72	.80	.75

TABLE 4.40

Oblique Multiple-Group Factor (Structure) Analysis for the 3 to 6 Years Subsample

Variables	I	II	I	II	III	I	II	III	IV
Vocabulary	.815	.395	.832	.351	.506	.899	.324	.351	.506
Comprehension	.788	.373	.812	.269	.537	.899	.278	[illegible]	.537
Pattern Analysis	.457	.802	.460	.850	.435	.315	.507	.850	.435
Copying	.389	.769	.365	.850	.397	.272	.362	.850	.397
Quantitative	.646	.577	.687	.511	.492	.335	1.000	[illegible]	.492
Bead Memory	.545	.759	.518	.452	.851	.408	.475	.452	.851
Memory for Sentences	.791	.470	.602	.380	.851	.580	.363	[illegible]	.851
Sum of Squares	2.986	2.660	2.801	2.252	2.582	2.403	1.928	2.252	2.582

TABLE 4.41

Oblique Multiple-Group Factor Intercorrelations Analysis for the 3 to 6 Years Subsample

Factors	I	II	I	II	III	I	II	III	IV
I	1.000		1.000			1.000			
II	.597	1.000	.485	1.000		.353	1.000		
III			.658	.489	1.000	.345	.511	1.000	
IV						.580	.492	.489	1.000

The above results are supportive of the theoretical areas of ability as outlined by Thorndike, Hagen and Sattler (1986(a)(b)). Comparisons can also be made to the standardization sample data for this age group that is outlined in Table 6.3 of the *Technical Manual'*. It is evident that the oblique multiple groups Clinic sample data is relatively more supportive of the SBIV theoretical model than the standardization data. Unlike the standardization data the Clinic sample variable loadings are consistently congruent with the proposed SBIV model. Moderate correlations are observed between factors and this is also consistent with the SBIV model.

<u>Oblique Multiple-Group Analysis for the 7 to 11 Years Subsample</u>. Tables [illegible] to 4.44 display the two, three and four oblique multiple-group solutions f[illegible] 7 to 11 years subsample. Crystallized and Fluid abilities are clearly identified by the two component solution Congruence coefficients of .88 and .83 are obtained.

Similar to the youngest subsample the three component solution for the middle group provides support for the two component results. In addition to the previously identified Crystallized and Fluid abilities a Short-term memory component is delineated. Congruence coefficients are, however, much more moderate ranging from .60 to .74.

The four component solution for this age group is highly similar to that obtained for the youngest subsample. The first component is clearly Verbal and is supported by salient loadings from the three Verbal Reasoning tests used in this analysis. Like the youngest subsample the second component is defined by a single salient loading. This component is tentatively identified as a Quantitative factor. Again, consideration is given to the fact that this is the single Quantitative Reasoning variable included in this analysis. The second and third components are respectively identified as Abstract/visual and Short-term Memory. Congruence coefficient are moderate and range from .52 to .71. Correlations between factors are similarly moderate and range from .65.

Comparisons can be made to the standardization sample data for this age group that is displayed in Table 6.4 of the *Technical Manual'*. Verbal and Memory factors are

TABLE 4.42

Oblique Multiple-Group Factor (Pattern) Analysis for the 7 to 11 Years Subsample

Variables	I	II	I	II	III	I	II	III	IV
Vocabulary	.973	-.135	.901	-.155	.107	.899	-.011	-.113	.115
Comprehension	.896	-.057	.937	-.052	-.022	.827	.143	-.032	-.018
Absurdities	.617	.169	.832	.088	-.109	.915	-.132	.145	-.097
Pattern Analysis	.061	.815	.086	.705	.149	.076	.012	.707	.149
Copying	-.184	.888	-.086	1.032	-.149	-.076	-.012	1.030	-.149
Quantitative	.724	.089	.705	.119	.023	.000	1.000	.000	.000
Bead Memory	-.008	.804	-.116	.155	.751	-.054	-.088	.163	.752
Memory for Sentences	.880	-.051	.116	-.188	.902	.107	.011	-.184	.903
Memory for Digits	.811	-.016	-.027	-.097	.926	-.111	.121	-.117	-.923
Memory for Objects	.131	.725	.027	.131	.679	.058	-.044	.138	.681
Sum of Squares	4.141	2.664	2.801	2.252	2.582	1.654	1.031	1.459	1.458
Coefficients of Congruence	.85	.72	.74	.60	.71	.69	.52	.60	.71

TABLE 4.43

Oblique Multiple-Group Factor (Structure) Analysis for the 7 to 11 Years Subsample

Variables	I	II	I	II	III	I	II	III	IV
Vocabulary	.870	.610	.883	.524	.716	.909	.580	.524	.716
Comprehension	.853	.630	.885	.559	.690	.886	.647	.559	.690
Absurdities	.746	.641	.805	.568	.619	.845	.484	.568	.619
Pattern Analysis	.686	.862	.676	.868	.719	.640	.591	.868	.719
Copying	.496	.747	.485	.868	.517	.445	.463	.868	.517
Quantitative	.792	.643	.803	.607	.671	.648	1.000	.607	.671
Bead Memory	.608	.798	.587	.611	.768	.568	.481	.611	.768
Memory for Sentences	.841	.623	.711	.532	.861	.691	.574	.532	.861
Memory for Digits	.798	.604	.648	.544	.836	.602	.598	.544	.836
Memory for Objects	.674	.809	.657	.633	.794	.638	.534	.633	.794
Sum of Squares	5.553	4.941	5.249	4.138	5.267	4.924	3.757	4.138	5.267

TABLE 4.44

Oblique Multiple-Group Analysis Factor Intercorrelations for the 7 to 11 Years Subsample

Factors	I	II	I	II	III	I	II	III	IV
I	1.000		1.000			1.000			
II	.766	1.000	.699	1.000		.648	1.000		
III			.799	.712	1.000	.625	.607	1.000	
IV						.767	.671	.712	1.000

identified in addition to the General factor. Both Absurdities and Bead Memory fail to load on their respective factors. A fourth factor identified is a singlet. Therefore, it is evident that for the 7 to 11 years subsample the Clinic data more closely adheres to the SBIV theoretical model.

Oblique Multiple-Group Analysis for the 12 to 23 Years Subsample. Tables 4.45 to 4.47 display the two, three and four oblique multiple-group solutions for the 12 to 23 years subsample. Like the two younger subsamples Crystallized and Fluid abilities are identified in the two component oblique multiple-group solution. Congruence coefficients are .84 and .63 respectively. These two components are highly (.89) suggesting a strong relationship between them.

The three component solution is supportive of the two component solution. In addition to the previously identified Crystallized and Fluid abilities, a Short-term memory component is supported. Congruence coefficients are modest and range from .51 to .74.

The four component solution is highly similar to its younger subsample counterparts . The first component is identified as a Verbal factor. The second component is identified as a Quantitative factor. This latter variable is, however, defined by the two available quantitative variables. The remaining components are defined as Abstract/visual and Short-term memory respectively. Congruence coefficients are in the range of .51 to .64. Correlations are moderately high and range from .81 to .89.

Comparisons can be made to the standardization sample data that is displayed in Table 6.5 of the *Technical Manual'*. Loadings are higher for the Clinic subsample. The latter are also observed to more closely adhere to the cognitive areas of ability that are outlined by the test authors.

Oblique Multiple-Group Summation

The oblique-multiple group analysis findings are supportive of the four cognitive areas of ability outlined by Thorndike, Hagen and Sattier (1986(a)(b)). Relative to the standardization data, these results are more closely aligned to the SBIV theoretical model.

TABLE 4.45

Oblique Multiple-Group Factor (Pattern) Analysis for the 12 to 23 Years Subsample

Variables	I	II	I	II	III	I	II	III	IV
Vocabulary	.863	.029	.877	-.108	.131	[illegible]	-.098	-.056	.153
Comprehension	1.060	-.164	1.180	-.229	-.053	.878	.311	-.208	-.044
Absurdities	.773	.102	.824	.203	-.135	.972	-.213	.264	-.109
Pattern Analysis	-.020	.888	-.074	1.084	-.100	-.014	-.068	1.088	-.099
Copying	.263	.665	.074	.074	.100	.014	.068	.790	.099
Quantitative	1.011	-.135	.930	-.037	-.003	-.099	1.193	-.124	-.040
Number Series	.743	.217	.747	.170	.060	[illegible]	.744	.124	.040
Bead Memory	-.075	.942	-.052	.289	.654	-.063	.015	.280	.652
Memory for Sentences	.1.058	-.204	.397	-.190	.676	.335	.058	-.178	.681
Memory for Digits	.675	.155	-.135	.016	.987	-.278	.178	-.010	.976
Memory for Objects	.168	.924	-.211	-.110	1.085	.006	-.251	-.092	1.093
Sum of Squares	5.719	3.141	4.499	2.071	3.097	2.792	2.236	2.075	3.094
Coefficients of Congruence	.84	.63	.74	.51	.64	.60	.52	.51	.64

TABLE 4.46

Oblique Multiple-Group Factor (Structure) Analysis for the 12 to 23 Years Subsample

Variables	I	II	I	II	III	I	II	III	IV
Vocabulary	.888	.800	.896	.763	.797	.923	.783	.763	.797
Comprehension	.913	.783	.935	.758	.776	.936	.857	.758	.776
Absurdities	.865	.793	.885	.813	.735	.920	.762	.813	.735
Pattern Analysis	.774	.870	.786	.939	.709	.759	.762	.939	.709
Copying	.858	.900	.853	.939	.803	.819	.833	.939	.803
Quantitative	.891	.770	.896	.773	.766	.794	.969	.773	.766
Number Series	.938	.882	.948	.871	.839	.880	.969	.871	.839
Bead Memory	.768	.875	.757	.765	.838	.727	.739	.765	.838
Memory for Sentences	.876	.742	.811	.700	.864	.801	.760	.700	.864
Memory for Digits	.813	.758	.726	.692	.884	.677	.737	.692	.884
Memory for Objects	.658	.773	.625	.579	.816	.619	.582	.579	.816
Sum of Squares	7.027	7.309	7.654	6.829	7.110	7.238	7.084	6.829	7.110

TABLE 4.47

Oblique Multiple-Group Analysis Factor Intercorrelations for the 12 to 23 Years Subsample

Factors	I	II	I	II	III	I	II	III	IV
I	1.000		1.000			1.000			
II	.894	1.000	.873	1.000		.864	1.000		
III			.858	.805	1.000	.840	.849	1.000	
IV						.830	.829	.805	1.000

The presence of moderate to strong positive correlations between factors is also consistent with the SBIV model. Higher order factor analysis is, however, required to more fully examine the relationship between factors, and possible levels within the SBIV theoretical structure.

Hierarchical Factor Analysis

As discussed in detail in Chapter 3, a pivotal debate existed in the 1930's "between advocates of oblique simple structure and advocates of the notion of a general factor (i.e., one on which all the variables have none zero regression)...." (McDonald, 1985, p. 104). Spearman saw oblique simple structure as a methodological artifact designed to avoid a general factor. Others saw *'g'* as a means of explaining the correlations that were found to exist between Thurstone's primary abilities (McDonald, 1985). Several authors including McNemar (1964), Gorsuch (1983) and McDonald (1985) have reconciled the diverse nature of findings by examining the scope of various intelligence measures and higher order factoring. Gorsuch (1983) has summarized the relationship between another aspect of generalizability and higher order factoring. He writes:

> Primary factors indicate areas of generalizability. More generalization can occur within a factor than across factors, but this does not eliminate generalization across factors. When factors are correlated, some generalization is possible. These areas of generalization across the primary factors form the higher-order factors.
>
> The essential difference between the primary factors and the higher-order factors is that the primary factors are concerned with narrow areas of generalization. In some analyses, the reduction in accuracy when going from primary to second-order factors will be small; in other studies it may be quite great. It depends on the data being analyzed (p. 24).

The matrix of correlations among primary or first order factors, if they are rotated obliquely, can be factored in the same way as the original correlation matrix. A number of extraction procedures are available and these are discussed in detail by Gorsuch (1983) and McDonald (1985). The extraction and factoring process can be repeated if it is necessary. This process is typically stopped when a single factor or uncorrelated factors result (Gorsuch, 1984).

Gorsuch (1983) emphasizes that critics of the higher order factoring process often stress that such factors are too far removed from the real world. He believes that this criticism is unfounded as higher-order factors are measured with the same degree of accuracy as the original factor. To ensure that one is not stepping too far from reality, the relationship between each level of higher-order factors and the original variables is often determined in addition to the relationship between levels of factors. The presence of a strong theoretical framework is also viewed as an additional safeguard in grounding one's factors to the real world (Gorsuch, 1984).

Gorsuch (1983) is correct in observing that oblique rotations are, by their very nature, suggestive of overlapping factors. Areas of overlap are potentially indicative of higher order factors. Extraction of higher order factors is often undertaken to gain a more complete understanding of one's data.

Several hierarchical arrangements are examined. These are displayed in Figure 4.0. The SBIV theoretical model, logical variants and earlier research findings served as selection criteria for these models. The number of model permutations and combinations is a barrier to succinct analysis. This difficulty is magnified threefold when each of the age groups are treated separately. An additional difficulty is that the number of variables for each groups differed. Therefore, only the most salient findings are discussed here and interested readers are directed to Appendix C for a presentation of these ancillary results.

Unrotated and Orthogonally Rotated Hierarchical Principal Components for the 3 to 6 Years Subsample. Table 4.48 displays the four factor hierarchical principal components for the youngest subsample. Original variable loadings on these four factors are displayed in Table 4.49. The presence of a strong General component is clear. The second factor is identified as an Abstract reasoning component. The third is bipolar with salient loadings on verbal as well as visual-spatial items. The fourth factor is a singlet.

The three and two factor unrotated hierarchical principal components are also supportive of a '*g*' component. A second component in both solutions is bipolar with

TABLE 4.48

SBIV Four Factor Hierarchical Principal Components

for the 3 to 6 Years Subsample

Factors	I	II	III	IV
I	.718	-.202	-.619	-.245
II	.485	.861	.061	-.143
III	.715	-.327	.538	-.304
IV	.826	-.047	.037	.561

TABLE 4.49

SBIV Original Variable Loadings on the Four Factor Hierarchical Solutions

for the 3 to 6 Years Subsample

Variables	I	II	III	IV
Vocabulary	.698	-.201	-.476	-.289
Comprehension	.671	-.081	-.557	.009
Pattern Analysis	.658	.309	.316	-.394
Copying	.631	-.442	.580	-.143
Quantitative	.669	.527	.172	-.138
Bead Memory	.728	.255	.120	.537
Memory for Sentences	.747	-.221	-.211	.220

FIGURE 4.0

Possible Hierarchical Structures for the SBIV

4		
4 - 1		
4 - 2		
4 - 3		
4 - 2 - 1	3	
4 - 3 - 1	3 - 1	
4 - 3 - 2	3 - 2	2
4 - 3 - 2 - 1	3 - 2 - 1	2 - 1

salient loadings on both verbal and abstract reasoning items. The third component in the three component solution is likely a Crystallized ability. These results are displayed in Tables 4.50 to 4.52. The most outstanding feature of these solutions is the strong support for a *'g'* component.

The results from the three second order orthogonal Procrustes solutions are displayed in Tables 4.53 to 4.55. Both four-two and four-three second order component solutions support the presence of a General factor. An Abstract reasoning component is also found in both of these solutions. The third component in the latter solution is identified as a Verbal factor. It is the three-two orthogonal Procrustes solution where verbal-performance distinction is most readily apparent.

Unrotated and Orthogonally Rotated Hierarchical Principal Components for the 7 to 11 Years Subsample. Table 4.51 displays the four factor hierarchical principal components for the middle subsample. Original variable loadings on these four factors are displayed in Table 4.57. Similar to the younger subsample a strong General factor is present. The second component is likely a Crystallized ability. The third component is a singlet and the fourth is defined by two salient short-term memory loadings. Therefore, the General component is the only constant in comparison to the younger subsample.

TABLE 4.50

SBIV Three Factor Hierarchical Principal Components Solution

for the 3 to 6 Years Subsample

Factors	I	II	III
I	.889	.940	.281
II	-.322	.940	-.114
III	-.953	.019	.301

TABLE 4.51

SBIV Original Variable Loadings on the Three Factor Hierarchical Solutions

for the 3 to 6 Years Subsample

		Factors	
Variables	I	II	III
Vocabulary	.425	.540	.540
Comprehension	.348	.350	.718
Pattern Analysis	-.444	.657	.187
Copying	-.074	.921	-.199
Quantitative	-.553	.452	.451
Bead Memory	-.409	.360	.638
Memory for Sentences	.201	.522	.566

TABLE 4.52

SBIV (First Order) Two Factor Hierarchical Solutions

for the 3 to 6 Years Subsample

Variables	I	II
Vocabulary	-.658	-.537
Comprehension	-.622	-.608
Pattern Analysis	-.717	.387
Copying	-.639	.345
Quantitative	-.723	.293
Bead Memory	-.728	.056
Memory for Sentences	-.702	-.372
Sum of Squares	3.288	1.154
Proportion of Total Variance	.470	.165
Cumulative Variance	.470	.635

TABLE 4.53

SBIV Second Order (Four to) Three Factor Orthogonal Procrustes Solutions

for the 3 to 6 Years Subsample

	Factors		
Variables	I	II	III
Vocabulary	.598	-.087	.623
Comprehension	.687	-.148	.523
Pattern Analysis	.449	.651	.056
Copying	-.112	.734	.615
Quantitative	.647	.571	-.094
Bead Memory	.565	.512	.170
Memory for Sentences	.486	.160	.624
Coefficients of Congruence	.785	.773	.463

TABLE 4.54

SBIV Second Order (Four to) Two Factor Orthogonal Procrustes Solutions

for the 3 to 6 Years Subsample

Variables	I	II
Vocabulary	.689	.229
Comprehension	.599	.313
Pattern Analysis	.368	.627
Copying	.770	-.008
Quantitative	.254	.813
Bead Memory	.456	.622
Memory for Sentences	.741	.240
Sum of Squares	2.385	1.649
Proportion of Total Variance	.341	.236
Coefficients of Congruence	.74	.56

TABLE 4.55

SBIV Second Order (Three to) Two Factor Orthogonal Procrustes Solutions

for the 3 to 6 Years Subsample

Variables	I	II
Vocabulary	.687	.024
Comprehension	.492	-.040
Pattern Analysis	.215	.763
Copying	.656	.651
Quantitative	-.012	.714
Bead Memory	.010	.545
Memory for Sentences	.528	.184
Sum of Squares	1.470	1.848
Proportion of Total Variance	.210	.264
Coefficients of Congruence	.70	.83

TABLE 4.56

SBIV Four Factor Hierarchical Principal Components

for the 7 to 11 Years Subsample

Factors	I	II	III	IV
I	.859	-.341	.091	.370
II	-.268	-.950	-.082	-.138
III	.768	.066	-.622	-.139
IV	.845	-.014	.447	-.294

TABLE 4.57

SBIV Original Variable Loadings on the Four Factor Hierarchical Solutions

for the 7 to 11 Years Subsample

Variables	I	II	III	IV
Vocabulary	.639	-.568	.225	.244
Comprehension	.657	-.520	.204	.170
Absurdities	.800	-.210	-.036	.375
Pattern Analysis	.850	-.095	-.030	-.239
Copying	.499	-.410	-.645	-.286
Quantitative	.562	-.552	.032	-.194
Bead Memory	.858	.131	.184	-.323
Memory for Sentences	.560	-.551	.379	-.218
Memory for Digits	.454	-.635	.260	-.399
Memory for Objects	.703	-.354	-.152	-.036

Similar to the younger subsample the three and two factor unrotated hierarchical principal components are also supportive of a *'g'* component. A second component in both solutions is likely a Verbal factor. The third component is identified as an Abstract/visual factor and this is also similar to the youngest group.

The results from the three second order orthogonal Procrustes solutions for the middle subsample are displayed in Tables 4.58 to 4.63. Relative to the younger subsample, the overriding presence of a strong General component is much more clear. The majority of variables in each solution load in a salient fashion on all factors.

Unrotated and Orthogonally Rotated Hierarchical Principal Components for the 12 to 23 Years Subsample. Table 4.64 displays the four factor hierarchical principal components for the oldest subsample. Original variable loadings on these four factors are displayed in Table 4.65. Similar to the younger subsamples a strong General component is evident. This component is, however, relatively stronger. The remaining components are less well defined and include a singlet and two Short-term memory components. These latter components bear little resemblance to their younger sample counterparts.

The three and two factor unrotated hierarchical principal component solutions are also supportive of a General component. These results are presented in Tables 4.66 and 4.67. A Memory component is identifiable in the three factor solution. The identity of the remaining components is unclear.

The results from the second order orthogonal Procrustes solutions are displayed in Tables 4.68 to 4.71. The salient theme is a strong General factor. Both two component solution factors are general. Two factors are also General in the four-three solution. The third factor is largely Abstract/visual. These findings are highly consistent with the middle subsample results.

TABLE 4.58

SBIV Three Factor Hierarchical Principal Components Solution

for the 7 to 11 Years Subsample

Factors	I	II	III
I	.871	.298	.391
II	.889	.217	-.403
III	-.536	.843	-.034

TABLE 4.59

SBIV Original Variable Loadings on the Three Factor Hierarchical Solutions

for the 7 to 11 Years Subsample

		Factors	
Variables	I	II	III
Vocabulary	.274	.732	.430
Comprehension	.293	.728	.365
Absurdities	.669	.486	.282
Pattern Analysis	.578	.632	-.097
Copying	.293	.615	-.600
Quantitative	.138	.799	.004
Bead Memory	.608	.519	-.003
Memory for Sentences	.043	.843	.259
Memory for Digits	-.104	.891	.057
Memory for Objects	.433	.660	-.053

TABLE 4.60

SBIV (First Order) Two Factor Hierarchical Solutions

for the 7 to 11 Years Subsample

Variables	I	II
Vocabulary	.779	-.411
Comprehension	.784	-.346
Absurdities	.748	-.033
Pattern Analysis	.830	.203
Copying	.674	.489
Quantitative	.769	-.138
Bead Memory	.745	.168
Memory for Sentences	.765	-.409
Memory for Digits	.735	-.323
Memory for Objects	.786	.088
Sum of Squares	5.814	.896
Proportion of Total Variance	.581	.090
Cumulative Variance	.581	.671

TABLE 4.61

SBIV Second Order (Four to) Three Factor Orthogonal Procrustes Solutions

for the 7 to 11 Years Subsample

		Factors	
Variables	I	II	III
Vocabulary	.709	.077	.523
Comprehension	.666	.110	.536
Absurdities	.399	.427	.586
Pattern Analysis	.301	.463	.654
Copying	.515	.753	-.032
Quantitative	.673	.200	.358
Bead Memory	.085	.316	.825
Memory for Sentences	.675	-.091	.545
Memory for Digits	.729	-.058	.377
Memory for Objects	.514	.455	.414
Coefficients of Congruence	.687	.743	.651

TABLE 4.62

SBIV Second Order (Four to) Two Factor Orthogonal Procrustes

Solutions for the 7 to 11 Years Subsample

Variables	I	II
Vocabulary	.793	.321
Comprehension	.757	.358
Absurdities	.543	.624
Pattern Analysis	.463	.720
Copying	.589	.265
Quantitative	.744	.258
Bead Memory	.264	.827
Memory for Sentences	.743	.257
Memory for Digits	.770	.125
Memory for Objects	.629	.473
Sum of Squares	4.221	2.265
Proportion of Total Variance	.422	.227
Cumulative Variance	.422	.649
Coefficients of Congruence	.86	.76

TABLE 4.63

SBIV Second Order (Three to) Two Factor Hierarchical Orthogonal Solutions

for the 7 to 11 Years Subsample

Variables	I	II
Vocabulary	.674	.395
Comprehension	.667	.414
Absurdities	.364	.742
Pattern Analysis	.524	.678
Copying	.556	.394
Quantitative	.764	.273
Bead Memory	.407	.688
Memory for Sentences	.823	.187
Memory for Digits	.895	.051
Memory for Objects	.576	.540
Sum of Squares	4.174	2.371
Proportion of Total Variance	.417	.237
Cumulative Variance	.417	.654
Coefficients of Congruence	.84	.75

TABLE 4.64

SBIV Four Factor Hierarchical Principal Components

for the 12 to 23 Years Subsample

Factors	I	II	III	IV
I	.872	-.165	-.317	-.335
II	.720	.690	.067	.002
III	.779	-.249	.575	-.031
IV	.882	-.181	-.249	.356

TABLE 4.65

SBIV Original Variable Loadings on the Four Factor Hierarchical Solutions

for the 12 to 23 Years Subsample

Variables	I	II	III	IV
Vocabulary	.878	-.018	-.188	-.169
Comprehension	.885	.128	-.166	-.217
Absurdities	.870	-.092	.023	-.307
Pattern Analysis	.848	-.162	.356	-.112
Matrices	.899	.080	.144	.014
Quantitative	.859	.348	-.028	.004
Number Series	.942	.124	.053	-.037
Bead Memory	.833	-.014	.270	.228
Memory for Sentences	.833	.158	-.358	.008
Memory for Digits	.797	.187	-.090	.458
Memory for Objects	.712	-.487	-.340	.315

TABLE 4.66

SBIV Three Factor Hierarchical Principal Components Solution

for the 12 to 23 Years Subsample

Factors	I	II	III
I	.902	.232	.365
II	.815	-.579	-.029
III	.892	.295	-.343

TABLE 4.67

SBIV Original Variable Loadings on Three Factor Hierarchical Solutions

for the 12 to 23 Years Subsample

		Factors	
Variables	I	II	III
Vocabulary	.855	.206	.174
Comprehension	.858	.204	.298
Absurdities	.853	-.029	.244
Pattern Analysis	.861	-.346	.047
Matrices	.903	-.082	.087
Quantitative	.849	.128	.247
Number Series	.937	.012	.155
Bead Memory	.858	-.200	-.143
Memory for Sentences	.809	.418	.128
Memory for Digits	.812	.209	-.218
Memory for Objects	.718	.322	-.503

TABLE 4.68

SBIV (First Order) Two Factor Hierarchical Solutions

for the 12 to 23 Years Subsample

Variables	I	II
Vocabulary	.248	.844
Comprehension	.348	.828
Absurdities	.434	.774
Pattern Analysis	.459	.832
Matrices	.354	.832
Quantitative	.348	.808
Number Series	.364	.878
Bead Memory	.227	.793
Memory for Sentences	.081	.859
Memory for Digits	-.073	.862
Memory for Objects	-.385	.842
Sum of Squares	1.172	7.468
Proportion of Total Variance	.107	.679
Cumulative Variance	.107	.786

TABLE 4.69

SBIV Second Order (Four to) Three Factor Orthogonal Procrustes Solutions

for the 12 to 23 Years Subsample

Variables	Factors I	II	III
Vocabulary	.653	.168	.593
Comprehension	.754	.141	.489
Absurdities	.585	.373	.534
Pattern Analysis	.500	.677	.404
Matrices	.713	.432	.375
Quantitative	.875	.178	.259
Number Series	.780	.353	.415
Bead Memory	.593	.547	.339
Memory for Sentences	.747	-.054	.534
Memory for Digits	.722	.155	.363
Memory for Objects	.227	.131	.889
Sum of Squares	4.948	1.332	2.742
Proportion of Total Variance	.450	.121	.249
Cumulative Variance	.450	.571	.820
Coefficients of Congruence	.73	.68	.64

TABLE 4.70

SBIV Second Order (Four to) Two Factor Orthogonal Procrustes Solutions

for the 12 to 23 Years Subsample

Variables	I	II
Vocabulary	.736	.478
Comprehension	.819	.358
Absurdities	.691	.537
Pattern Analysis	.635	.585
Matrices	.806	.406
Quantitative	.914	.158
Number Series	.866	.391
Bead Memory	.700	.451
Memory for Sentences	.791	.305
Memory for Digits	.775	.262
Memory for Objects	.349	.789
Sum of Squares	6.164	2.319
Proportion of Total Variance	.560	.211
Cumulative Variance	.560	.771
Coefficients of Congruence	.85	.73

TABLE 4.71

SBIV Second Order (Three to) Two Factor Hierarchical Orthogonal Solutions

for the 12 to 23 Years Subsample

Variables	I	II
Vocabulary	.815	.330
Comprehension	.817	.333
Absurdities	.677	.519
Pattern Analysis	.499	.782
Matrices	.687	.592
Quantitative	.765	.390
Number Series	.769	.535
Bead Memory	.582	.661
Memory for Sentences	.901	.130
Memory for Digits	.782	.302
Memory for Objects	.772	.156
Sum of Squares	6.047	2.460
Proportion of Total Variance	.550	.224
Cumulative Variance	.550	.774
Coefficients of Congruence	.85	.70

Summation for the Unrotated and Orthogonally Rotated Hierarchical Principal Components

The unrotated hierarchical principal component results are supportive of a General factor. Support for additional group factors is more consistent in the younger subsamples. Verbal and Abstract/visual factors are evident. Little support is found for such group factors in the oldest subsample.

The orthogonal hierarchical models are also supportive of a strong General component. The first and third components in the second order four-three solutions are largely general. The second component is typically identified as being Abstract/visual in nature. However, outside variables are also observed to load on this factor. Consequently, the explanatory power of this solution is questionable.

The four-two orthogonal solution is a relatively better option. A *'g' factor* is found across all three subsamples. The second component in the youngest subsample is largely an Abstract/visual factor. Memory, quantitative and verbal variables are, however, observed to load on this second factor. For the oldest subsample, this component is identified as a more moderate *'g'* factor. Overall, the four-component solution is perceived to a relatively better explanation of the data.

Oblique Hierarchical Analysis

Oblique Hierarchical Analysis for the 3 to 6 Years Subsample. Figure 4.0 outlines all of the hierarchical models that are examined. For reasons of clarity and parsimony much of these results have been placed in Appendix C. This is to ensure that the mere bulk of the analyses does not detract from the task at hand.

The four-three model is unsatisfactory for the youngest subsample (See Appendix A.0 to A.6). All loadings are greater than one. This is possible because they are regression weights. Therefore, little understanding is gained from this model for the 3 to 6 years subsample.

The four-three-one model is similarly unacceptable (See Appendix A.6). The highest component is bipolar. A verbal-performance split is readily apparent. This factor adds little to the understanding of the test battery. It is also inconsistent with the SBIV theoretical model which proposes an overriding General factor.

The majority of the loadings on both factors in the four-three-two model are salients (See Appendix A.3 to A.5). This is expected as the correlations between factors at the second level are high. Relatively stronger loadings are obtained from Verbal Reasoning variables on the first factor. Abstract Reasoning variable loadings are stronger on the second. Other variables line up in anticipated directions. Copying is the least distinctive variable.

The four-three-two-one model results are, as expected, of little practical value (See Appendix A.7). Negative salient loadings are observed on Bead Memory, Pattern Analysis, and Quantitative variables. An Abstract Reasoning component is suggested. A single positive salient loading is obtained from the Vocabulary test. These results are largely inconsistent with the proposed SBIV model.

Both factors in the four-two oblique model are General components (See Appendix A.8 to A.10). These results are similar to the third level of the four-three-two model discussed earlier. As is anticipated the four-two-one is of little value (See Appendix A.11) for this age group.

The four-one oblique hierarchical solution is displayed in Table 4.72. All variables have moderately strong salient loadings on this factor. This result is consistent with the unrotated principal components solution. Both solutions are supportive of a single General factor.

Findings relatively more consistent with the SBIV model are found in the three-two oblique hierarchical model. At the second level a clear distinction between Verbal and Abstract Reasoning components is evident. Both Bead Memory and Quantitative are aligned with the later factor. This is a relatively common finding throughout this analysis. These results are outlined in Tables 4.73 to 4.75.

The three-one and three-two-one hierarchical models are remarkably consistent (See Appendix A.12 and A.13). Both components are bipolar. Positive salients are observed on the two verbal tests. Negative salients are observed on Bead Memory, Pattern Analysis and Quantitative tests. These results are consistent with four-three-one model and offer little to the understanding of the overall make-up of the SBIV.

The two-one oblique factor model is the final model examined for each group. These results are similar to the four-one model and are displayed in Table 4.76. This solution is supportive of a General component.

In summary, relatively more support is found for a modified SBIV model in this youngest group. More useful explanations are offered by the four-one, three-two and two-one models. These results suggest that the SBIV theoretical structure is equivocal and only partially supported in this age group.

Oblique Hierarchical Analysis for the 7 to 11 Years Subsample. Like the youngest subsample, the four-three model for the middle subsample is largely unsatisfactory (See Appendix A.14 to A.16). The first and third factors are General components. The middle factor is identified as an Abstract/visual component. Noteworthy, is that both Absurdities and Memory for Objects load on this factor. This is a common finding for this age group.

The four-three-one model is similarly unacceptable (See Appendix A.17). The highest factor is likely an Abstract/visual component. Like the previous solution, Absurdities, Pattern Analysis and Bead Memory are joint salients. The interpretative value of this model is limited.

TABLE 4.72

SBIV Four to One Hierarchical Model

for the 3 to 6 Years Subsample

Variables	I
Vocabulary	.698
Comprehension	.671
Pattern Analysis	.658
Copying	.631
Quantitative	.669
Bead Memory	.728
Memory for Sentences	.747

TABLE 4.73

SBIV Three to Two Oblique Factor (Pattern) Hierarchical Model

for the 3 to 6 Years Subsample

Variables	II	III
Vocabulary	.699	.113
Comprehension	.498	.023
Pattern Analysis	.261	.798
Copying	.702	.742
Quantitative	.028	.718
Bead Memory	.041	.551
Memory for Sentences	.547	.254
Congruence Coefficients	.700	.836

TABLE 4.74

SBIV Three to Two Oblique Factor (Structure) Hierarchical Model for the 3 to 6 Years Subsample

Variables	I	II
Vocabulary	.679	-.014
Comprehension	.493	-.067
Pattern Analysis	.116	.750
Copying	.568	.614
Quantitative	-.103	.713
Bead Memory	-.059	.543
Memory for Sentences	.501	.155

TABLE 4.75

SBIV (Three to) Two Correlations Between Factors for the 3 to 6 Years Subsample

Factors	I	II
I	1.000	
II	-.181	1.000

TABLE 4.76

SBIV (Two-to-One) Oblique Factor Hierarchical Model

for the 3 to 6 Years Subsample

Variables	I
Vocabulary	-.658
Comprehension	-.622
Pattern Analysis	-.717
Copying	-.639
Quantitative	-.723
Bead Memory	-.728
Memory for Sentences	-.702

The four-three-two and four-two models are much the same. This is somewhat surprising given the four-three model findings. Both of the former solutions are supportive of Crystallized and Fluid ability interpretations. These results are displayed in Tables 4.77 to 4.82. Exceptions to predicted patterns include the loading of Absurdities on the Fluid component and the somewhat similar loading by Memory for Objects on both factors.

The four-three-two-one, four-two-one and four-one models are similar at the highest level. All three solutions are supportive of a General component at the highest level. The four-one model is, however, more consistent with the unrotated principal components and is displayed in Table 4.83 (also See Appendix A.18 to A.20). It is preferred as it avoids the complexities of the intermediary levels.

The three-two oblique solution for the middle group is parallel to the younger subsample. The majority of variables on both factors are salients. However, when the relative strength of loadings are taken into account a Crystallized-Fluid distinction is evident. Consistent with earlier findings Absurdities is salient on the Fluid component. Also, Copying and Memory for Objects have much the same salient loadings on both factors. These results are displayed in Tables 4.84 to 4.86.

The four-one, three-two-one and two-one models are all supportive of a General component at the highest level. Results for the four one model are displayed in Table 4.80. The four-one model is supported by the unrotated hierarchical components results.

Unlike the above single component solutions, the three-one model is not supportive of a General factor (See Appendix A.21 and A.22). An Abstract/visual component is identified. Similar to earlier results this factor is identified by salient loadings from Memory for Objects, Pattern Analysis, Bead Memory, and Absurdities.

TABLE 4.77

SBIV Four-Three-Two Oblique Factor (Pattern) Hierarchical Model

for the 7 to 11 Years Subsample

Variables	I	II
Vocabulary	.492	.222
Comprehension	.459	.267
Absurdities	.295	.602
Pattern Analysis	.191	.705
Copying	.728	.353
Quantitative	.558	.200
Bead Memory	-.116	.787
Memory for Sentences	.399	.131
Memory for Digits	.519	.018
Memory for Objects	.472	.464
Congruence Coefficients	.770	.817

TABLE 4.78

SBIV Four-Three-Two Oblique Factor (Structure) Hierarchical Model

for the 7 to 11 Years Subsample

Variables	I	II
Vocabulary	.554	.360
Comprehension	.533	.395
Absurdities	.463	.684
Pattern Analysis	.387	.758
Copying	.826	.556
Quantitative	.614	.356
Bead Memory	.103	.754
Memory for Sentences	.436	.242
Memory for Digits	.524	.163
Memory for Objects	.601	.596

TABLE 4.79

SBIV (Four-Three-) Two Oblique Factor Intercorrelations

for the 7 to 11 Years Subsample

Factors	I	II
I	1.000	
II	.279	1.000

TABLE 4.80

SBIV (Four-) Two Oblique Factor (Pattern) Hierarchical Solution

for the 7 to 11 Years Subsample

Variables	I	II
Vocabulary	.736	.197
Comprehension	.681	.247
Absurdities	.335	.601
Pattern Analysis	.207	.729
Copying	.535	.177
Quantitative	.708	.135
Bead Memory	-.056	.896
Memory for Sentences	.707	.135
Memory for Digits	.792	-.024
Memory for Objects	.492	.408
Congruence Coefficients	.887	.775

TABLE 4.81

SBIV (Four-) Two Oblique Factor (Structure) Hierarchical Solution for the 7 to 11 Years Subsample

Variables	I	II
Vocabulary	.838	.582
Comprehension	.811	.604
Absurdities	.650	.776
Pattern Analysis	.589	.837
Copying	.628	.457
Quantitative	.779	.506
Bead Memory	.413	.867
Memory for Sentences	.778	.505
Memory for Digits	.780	.391
Memory for Objects	.706	.665

TABLE 4.82

SBIV (Four-) Two Oblique Factor Intercorrelations for the 7 to 11 Years Subsample

Factors	I	II
I	1.000	
II	.524	1.000

TABLE 4.83

SBIV Three to Two Oblique Factor (Pattern) Hierarchical Model for the 7 to 11 Years Subsample

Variables	I	II
Vocabulary	.590	.333
Comprehension	.578	.353
Absurdities	.167	.748
Pattern Analysis	.352	.657
Copying	.467	.347
Quantitative	.719	.188
Bead Memory	.227	.684
Memory for Sentences	.806	.088
Memory for Digits	.921	-.068
Memory for Objects	.446	.501
Congruence Coefficients	.847	.753

TABLE 4.84

SBIV Three to Two Oblique Factor (Structure) Hierarchical Model for the 7 to 11 Years Subsample

Variables	I	II
Vocabulary	.719	.560
Comprehension	.714	.576
Absurdities	.455	.813
Pattern Analysis	.605	.793
Copying	.601	.528
Quantitative	.792	.466
Bead Memory	.491	.771
Memory for Sentences	.840	.399
Memory for Digits	.894	.287
Memory for Objects	.640	.674

TABLE 4.85

SBIV (Three to) Two Correlations Between Factors

for the 7 to 11 Years Subsample

Factors	I	II
I	1.000	
II	.386	1.000

TABLE 4.86

SBIV Four to One Hierarchical Model

for the 7 to 11 Years Subsample

Variables	I
Vocabulary	.639
Comprehension	.657
Absurdities	.800
Pattern Analysis	.850
Copying	.499
Quantitative	.562
Bead Memory	.858
Memory for Sentences	.560
Memory for Digits	.454
Memory for Objects	.703

In summary, like the youngest subsample, only modest support is found for the SBIV theoretical structure. More useful information is offered by the four-one and four-two models. Bead Memory and Absurdities are observed to consistently load on Abstract/visual or Fluid components. Memory for Objects is observed to frequently load on more than one factor in a relatively similar fashion. These results are generally consistent with younger subsample results.

Oblique Hierarchical Analysis for the 12 to 23 Years Subsample. The four-three model results for this group are relatively more clear than those obtained for the younger subsamples (See Appendix A.23 to A.25). The first component is largely general. The second factor is identified by salient loadings from Absurdities, Matrices, Bead Memory and Pattern Analysis. This pattern is consistent with earlier results. It is identified as an Abstract reasoning component. The third component is complex and is defined by salient loadings on both verbal and memory (excluding Bead Memory) tests. Therefore, the utility of this model and its alignment with the proposed SBIV model is still questionable despite the relatively more clear pattern present.

The four-three-two and four-two models are quite similar at the highest level (See Appendix A.26 to A.31). The first component in both solutions is a General factor. The second component is identified as an Abstract/visual factor, however, Vocabulary loads salients in both solutions. These solutions are, therefore, viewed as being relatively unsatisfactory for interpretative purposes.

The four-three-one, four-three-two-one, four-two-one and four-one hierarchical solutions are consistent in identifying a General factor at the highest level (See Appendix A.32 to A.34). The results of the four-one solution are displayed in Table 4.87. This solution is preferred as it is corroborated by the unrotated hierarchical components. It also avoids interpretative difficulties encountered by the more complex models.

TABLE 4.87

SBIV Four to One Hierarchical Model

for the 12 to 23 Years Subsample

Variables	I
Vocabulary	.878
Comprehension	.885
Absurdities	.870
Pattern Analysis	.848
Matrices	.899
Quantitative	.859
Number Series	.942
Bead Memory	.833
Memory for Sentences	.833
Memory for Digits	.797
Memory for Objects	.712

The three-two oblique solution is largely unsatisfactory for the oldest group (See Appendix A.35 to A.37). Although loadings are generally in the predicted direction for a Fluid and Crystallized abilities distinction, many variables load on both factors. Memory for Objects, however is observed to load only on the Crystallized factor. Absurdities is seen to have a stronger loading on the Fluid component. Therefore, these latter variable loadings are inconsistent with the hypothesized SBIV model. For these reasons this model is deemed largely unsatisfactory.

A general component is supported by both the three-two-one and three-one models (See Appendix A.38 and A.39). The difficulties both of these models encounter are at their lower levels. However, the three-one model did receive some support from the orthogonal Procrustes model. The two-one model is least useful. Salient loadings are observed on Verbal, Abstract/visual and Quantitative reasoning variables. However, Vocabulary and three of the memory variables are left undefined.

Summation for the Oblique Hierarchical Analysis

In summation, the four-one models likely provide the best explanation for the 12 to 23 year data. These models avoid the complexities at intermediate levels which are not, in themselves, empirically clear. They also receive some support from the theoretical constructs underlying the SBIV model.

Summation for the Hierarchical Analysis

The results of the hierarchical analysis are complex and largely equivocal. Four, three and two unrotated hierarchical components are extracted for each age group. A large *'g'* component is found in each solution at every age. The presence of a strong *'g'* component is clearly the most consistent finding. The remaining unrotated hierarchical principal component results are less promising.

The four unrotated hierarchical principal component solutions are supportive of a *'g'* component. Singlets are also found at each age level. A single Short-term memory component is found at the 7 to 11 years level. Two such factors are identified at the highest level under this solution.

The remaining factors are unique to each age group. At the youngest level an Abstract reasoning component is identified. So too is a bipolar Verbal-Visual/spatial component. At the 7 to 11 years subsample a Crystallized ability is found.

The three unrotated hierarchical component solutions are somewhat similar to the four component solutions. The presence of *'g'* is the single common thread between age groups. Again, group factors are more clearly evident in the younger groups.

A bipolar Verbal-Visual/spatial component is found in the youngest subsample. This factor appears to divide into two components at the 7 to 11 years level. Unlike the older age groups a Crystallized ability is found in the youngest age group. It is only at the oldest age group that a Short-term memory factor is found to emerge in addition to the Crystallized/Fluid split.

The two component unrotated hierarchical principal component solution results are parallel to the three and four component findings. A strong *'g'* component that increases in influence with age is the single most consistent finding. Similar to its multiple component counterparts a bipolar Verbal-Abstract/visual component is found at the youngest age level. The second factor at the 7 to 11 years is largely a Verbal component. This factor at the highest age level is largely undefinable.

In brief, there are three major findings from the unrotated hierarchical principal components analysis. A strong *'g'* component is found within all three solutions at each age level. The magnitude of this component is found to increase with age.

The second finding is that within age group findings appear to be more consistent than inter-age group results. Verbal and Visual/spatial factors receive some support in the youngest age group. Moderate support for these group factors is found at the 7 to 11 years level. Consistent group factors are not found for the oldest group under these multiple level hierarchical solutions. Therefore, the equivocal nature of the four cognitive areas of ability in the unrotated hierarchical analysis is the third major finding.

Several orthogonally rotated hierarchical principal component models are also examined. The four-three models under orthogonal rotation are largely unsatisfactory for interpretation. In addition to *'g'* both Verbal and Abstract/visual components are found at the youngest age level. For the middle group two of three variables are largely *'g'*. Within each of these factors relatively higher loadings are found on either Verbal or Abstract/visual variables. The third component is largely Abstract/visual. Relatively similar results are obtained for the oldest group.

The four-two model under orthogonal rotation is also unsatisfactory. Both factors are largely *'g'* for the two oldest groups. Relatively higher loadings do divide between the two factors along Verbal-Abstract/visual lines. At the youngest age level an Abstract/visual component is found in addition to the *'g'* component. However, verbal, memory and quantitative tests are observed to load on this variable.

A more clear distinction between Verbal and Nonverbal reasoning components is found under the three-two orthogonal model. At both the youngest and oldest ages this distinction is present. However, the strength of *'g'* factor is more readily evident in the older group and pervades both components.

The middle age group findings are less clear. The first component is largely general. The second is bipolar, dividing between Abstract/visual and Verbal themes. Given the factor discrepancies between age groups and the strength of *'g'* in both factors at the 12 to 23 years level this model is viewed as possessing relatively little explanatory value. Overall, the unrotated hierarchical component models are most noticeably supportive of a strong *'g'* factor interpretation. Less and inconsistent support is found for multiple-factor interpretations under orthogonal rotations.

Several oblique hierarchical models are then evaluated. The four-three model is largely unsatisfactory. At the youngest level all variables are salients and the factors are highly correlated. It is possible that the factors are measuring the same thing. At the 7 to 11 ages two *'g'* components are present. The third is likely Abstract/visual.

The three components at the highest age group are identified as *'g'* Abstract/visual and a blended factor consisting of both Verbal and Short-term memory variables. This result is somewhat consistent with orthogonal Procrustes analysis. It is also moderately consistent with the SBIV model as the four cognitive areas of ability are seen to surface.

The four-three-two model is disappointing. This is not surprising given the earlier findings. For the youngest age group Verbal and Abstract reasoning factors are identified. These factors are more readily described as Crystallized and Fluid abilities in the 7 to 11 age group. In this oldest group an Abstract/visual component is found in addition to the *'g'* component. However, loadings from outside the Abstract/visual component sphere are also found. It is evident then, that even under this model the broad influence of *'g'* increases as subjects age.

The above trend is clearly demonstrated in both the four-three-one and four-three-two-one oblique models. In the youngest subsample both solutions have identified an Abstract reasoning component at the highest level. General factors of increasing strength are found at the highest level for both solutions in the older groups.

The results of the four-two analysis are somewhat of an anomaly. At the youngest level both factors are General. For the two older groups a single General component is found along with an Abstract/visual reasoning factor. The latter is defined by Absurdities, Pattern Analysis, Bead Memory, and Memory for Objects. Copying is noticeably absent. These results are inconsistent with the previously identified trend of increasing influence of *'g'* as subjects age.

The four-two-one results are largely an extension of the four-two model. At the youngest age level an Abstract/visual component is found. For the two older groups a General component is found.

The four-one model is observed to be consistent across all age groups. All variables are salients indicating a strong General component. This component is strongest for the 12 to 23 years age group.

There are three main advantages of this model. The first is parsimony. This model provides a clear understanding of the data with the minimal complexities. Secondly, it provides an explanation that is consistent for each age group. Therefore, despite its simplicity the four-one model possesses the widest explanatory power. Thirdly, it is consistent with the test construction design. Therefore it does not require any practical SBIV modifications and only slight theoretical revisions.

At the highest level the three-two model is an extension of its primary components. Verbal and Abstract/visual components are identified in the youngest group. These factors are seen to broaden to Crystallized and Fluid ability at the two older levels. However, the lack of specificity in some variables is a barrier to interpretation.

Some additional support for the three factor interpretations is garnished when the model is extended to its ultimate form. At the highest level of the three-two-one model a strong *'g'* factor is found for the two oldest groups. A bipolar, Verbal-Abstract reasoning component is found at the highest level for the youngest group. These findings are remarkably similar to the three-one model that is criticized for its lack of ability to provide a consistent explanation across age groups.

The two-one model is the final hierarchical model examined. The strong presence of a *'g'* is found at the primary level. At the second level a *'g'* component is observed in the two younger groups. However, only six variables in the oldest group load in a single direction on this factor. Therefore, this model is unacceptable for the older group. The lack of specificity of variables such as Copying and Memory for Objects is also a weakness of this model at the primary level.

Overall, the results of the hierarchical analysis are less than promising. Consistent support for the hierarchical structure espoused by Thorndike, Hagen and Sattler

(1986(a)(b)) is not found. Neither is unchanging support for other multi-level models. Relatively greater support is found for a modified two-tier model. At the first level the four cognitive areas are found. At the second level *'g'* is identified as the unifying force. This model is observed to receive relatively more consistent support across all three age groups. Overall, these results suggest caution should be utilized in interpreting the SBIV data according to the existing hierarchical structure. Much research is still required.

<u>Summation</u>

The Clinic sample descriptive statistics are generally consistent with standardization sample results. Clinic sample intercorrelations , means and standard deviations are generally only slightly higher. Exceptions to this trend are generally confined to tests that reflect fewer subjects. Overall, the descriptive statistics for all three subsamples and the entire sample are compatible with the SBIV standardization data. These findings are supportive of the Clinical sample and lend credibility to subsequent analyses.

The unrotated principal component results are supportive of a strong General factor. A second finding from the unrotated principal component data is that not all variables are observed to load on the hypothesized factors (e.g., Bead Memory, Absurdities, and Memory for Sentences). Certain variables are also observed to load on different factors at various ages (e.g., Quantitative). Obtained loadings are often small. Other variables are observed to lack specificity under different solutions (e.g., Copying and Memory for Objects).

An orthogonal principal components analysis with varimax rotations are then employed on differing numbers of extracted factors. Modest support for three of the four SBIV cognitive areas of ability is found when varimax procedures are employed These three components are; Verbal, Nonverbal/visual and Short-term memory. These components are, however, less well defined in the two older subsamples. A separate Quantitative factor is not found and this is consistent with Sattler's (1988) varimax analysis of the standardization data. Overall, the exploratory rotational results are perceived to better match the proposed SBIV model than the standardization data themselves. One possible explanation for this is the test authors' extraction of the influence of *'g'* in their standardization data. The influence of *'g'* is not directly extracted in the Clinic data.

Orthogonal Procrustean techniques are then employed and extracted components are rotated to the various hypothetical models proposed to fit the SBIV structure. Moderate support is found for Crystallized and Fluid abilities in the two younger subsamples when only two components are extracted. Less support is found for the two factor solution in the older group largely due to the strength of the General factor.

The three component solution is least supported. The hypothesized Fluid ability is identified as being somewhat weaker for the oldest subsample likely due to the strength of the General factor. Overall, the proposed Short-term memory component is the most tenuous. This is largely due to the findings that designated memory variables are found to load on more than one factor. This is also true for other components.

The four component orthogonal Procrustes solution is also complex and many of the two and three component findings are relevant. However, relatively more solid support is found for Verbal, Abstract/visual and Short-term memory factors. The Quantitative factor is the least well defined, however, it is likely restricted by the number of potential variables assigned to it. Overall, the four component solution is likely a closer approximation of the data.

Several ancillary orthogonal Procrustes analyses are conducted. Randomizations are employed to examine 5,000 possible alternative models. These results are supportive of the two factor model as the best explanation for the middle group data. The oldest group results are best explained by a three factor interpretation. The four factor model is best able to explain the youngest group results. It is, also the only model that reaches significance across all three age groups. Therefore, this model is viewed as possessing relatively greater explanatory power and as such is preferred.

Three additional criteria are proposed by Helmes (1989) to evaluate internal test structure. Helmes criteria are based on the number of items having their highest loadings on designated and nondesignated factors, and the mean loading of targeted variables. These results are remarkably consistent with the randomization data results and support earlier findings. These results are also clear in their warning that few generalizations should be made from individual variables given their potentially tenuous nature (e.g., Quantitative).

An oblique multiple group analysis is then employed. This technique is generally perceived as an excellent means for evaluating theoretical test structures. It is a method that is highly similar to that employed by the test authors. The oblique multiple group results are generally supportive of the various theoretical substructures that are outlined by the test authors. Moderate to strong support is found for the proposed two, three and four factor models and, in general, this analysis is the most supportive of the SBIV theoretical model.

Hierarchical analysis is used to clarify the relationships between test variables at various levels. These techniques are essential given the multi level theoretical model proposed by Thorndike, Hagen and Sattler (1986(a)(b)). A review of the SBIV theoretical structure and logical variants is provided. The results of the hierarchical analysis are complex and do not lend to succinct analysis. They suggest that the four-one model is likely the most acceptable interpretation of the data. Other arrangements identify various aspects of the SBIV model but fail to provide consistent explanations that are theoretically and/or empirically in agreement with the SBIV model substructures. Therefore, like the various group factor analyses the most consistent finding is the tentative presence of the four cognitive areas of ability and an overriding general ability. Therefore, only partial support of the SBIV is found using the various factor analytic procedures and a suggestion is present that various developmental factors may wish to be considered.

CHAPTER 5

DISCUSSION

Introduction

The identification of fundamental intellectual abilities has preoccupied philosophers and psychologists for two thousand years. Within the last century the accurate measurement of intelligence has been a primary activity of psychologists. Initially, the dearth of sophisticated statistical methods required to develop and evaluate psychometric tests was a significant barrier (Nunnally, 1978). However, by the turn of the century both the means of measurement and methods of analyses were introduced.

Theoretical Underpinnings of the SBIV

Binet and Simon (1905) were credited for developing the first practical test of human ability. This instrument is the direct ancestor of the SBIV and most other contemporary measures of human intelligence (Tuddenham, 1962). The early Binet, however, is largely atheoretical. Intelligence is viewed as consisting of a single fundamental faculty-judgement.

Since 1905 a number of intellectual theories have evolved. Weinberg (1989) suggested that at the broadest level intelligence theorists can be divided into two camps. These are "lumpers" and "splitters". Lumpers are individuals who emphasize broad, general abilities. Splitters are theorists who espouse multiple independent faculties (Mayr, 1982).

Binet and Spearman are correctly identified as "lumpers". Thurstone, Guilford and Gardner are cited as "splitters". Hierarchical theorists such as Horn, Cattell and Vernon are viewed as occupying middle ground and as offering a rapprochement between these two antithetical positions. The SBIV is an exemplar of such a reconciliation of diverse, mainstay views.

Empirical Underpinnings of the SBIV

The SBIV is also able to build on almost fifty years of factor analytic work on earlier editions. This research is supportive of a strong general factor (Burt and John, 1942(a)(b); Hallahan, Ball and Payner, 1973). Several group factors are also frequently found including: Verbal (Burt and John, 1942(a)(b); Hallahan, Ball and Payne, 1973; Jones, 1949, 1954; McNemar, 1942; Ramsey and Vane, 1970; Thompson, 1984; Wright, 1939), Nonverbal-Visual/spatial (Burt and John, 1942; Hallahan, Ball and Payne, 1973; Jones, 1949, 1954; McNemar, 1942; Ramsey and Vane, 1970; Thompson, 1984; Wright. 1939), Number, (Burt and John, 1942; Jones, 1949, 1954; McNemar, 1942; Wright, 1939), and Memory (Burt and John, 1942; Jones, 1949, 1954; McNemar, 1942; Ramsey and Vane, 1970; and Thompson, 1984).

Various explanations have been offered for differences in these factor analytic results. Gorsuch (1983) and Mulaik (1972) have addressed these reasons in detail. A widely held belief is that these differences are due to choice of rotational method. Helmes (1989) has emphasized that substantial differences in rotational methods have not been demonstrated.

SBIV Theoretical Structure

The SBIV is, therefore, able to rely on considerable historical, theoretical and empirical support. This led Thorndike, Hagen and Sattler (1986(a)(b)) four areas of cognitive abilities that are arranged into a three level hierarchical model. At the top level is 'g', a general reasoning factor. Three broad factors are found at the second level. These are Crystallized, Fluid-analytic and Short-term memory abilities. Three somewhat more specific factors are found at the third level. These are Verbal reasoning, Quantitative reasoning and Abstract/visual reasoning. The former two are aligned beneath the Crystallized ability. The latter variable is found under the Fluid-analytic ability.

SBIV Empirical Structure

The largest standardization sample of any individualized test to date was employed. An attempt was made to create a sample representative of the U.S. population. A weighting procedure was used in an attempt to correct any sample deficiencies.

Reliability data is provided by Thorndike, Hagen and Sattler (1986(a)(b)). The SBIV is found to possess good internal consistency reliabilities (Sandoval and Irvin, 1988; Vernon, 1987). Concurrent validity data is also supportive of the SBIV (Carvajal et al., 1987(a)(b); Clark, 1988; Fritzke, 1988; Rothlisberg, 1988; and Thorndike, Hagen and Sattler, 1986(b)). The SBIV is likewise found to be able to discriminate between various special groups (Clark, 1988; Fritzke, 1988; Hofman, 1988; and Meloff, 1987).

It is the SBIV factorial results that have been most seriously questioned (Osberg, 1986; Sandoval and Irvin, 1988; Slate, 1986; and Vernon, 1987). Consensus is obtained on the presence of a strong 'g' component (Osberg, 1986; Sandoval and Irvin, 1988; Sattler, 1988; Slate, 1986 and Vernon, 1987). These same authors reject Thorndike, Hagen and Sattler's (1986(b)) remaining interpretation of the unnamed variant of confirmatory factor analysis data. They emphasize that these results are not supportive of the four cognitive areas of ability. Vernon (1987) stresses that the total group factor content (variance) is only 11 percent. He also points out that many tests possess low relationships with other tests in their proposed hypothetical area in addition to their failure to load on designated factors.

Upon review of the SBIV technical data these criticisms appear to have some merit. The presence of a strong 'g' component is readily apparent. However, different numbers of group factors are identified at each age level analyzed.

At the youngest level (2 to 6 years) only two group factors are identified. These are Verbal and Abstract/visual. Importantly, Absurdities is observed not to load in a salient fashion on the Verbal factor whereas this is true for Memory for Sentences. Pattern Analysis is a nonsalient on the second factor. However, Bead Memory is observed to be the most significant salient on this second component. These two group factors are observed to account for only 14 percent of the total variance.

For the middle group (7 to 11 years) three group factors are identified by Thorndike, Hagen and Sattler (1986b). The first group component is identified as a Verbal factor. Again, Absurdities is not a salient on this factor. The second factor is described by three Short-term memory variable salients. Bead Memory is a nonsalient. This second component account for five percent of the total variance.

The third and final component is identified as Abstract/visual. It is, however, a singlet, defined solely by Pattern Analysis. Copying and Matrices are nonsalients. As such, this factor is tenuous by most standards. This third component accounts for 3.5 percent of the total variance. Overall, the three group factors identified at this age are observed to explain 12.5 percent of the variance.

Four group factors are identified in the technical data for oldest (12 to 23 years) group. The first component is a Verbal factor. It is defined by three verbal tests. This component is able to explain 4.8 percent of the variance.

The second component is identified as a Memory factor. Two of the memory variables administered are salients. Memory for Sentences is a near salient. Bead Memory is, however, a nonsalient. This Memory factor is able to explain 3.7 percent of the total variance.

The third component identified for the oldest group is Abstract/visual. Pattern Analysis and Paper Folding and Cutting define this factor. Importantly, Matrices is a nonsalient. This factor is able to account for 3.4 percent of the total variance.

Similar to the middle group the final factor identified is a singlet. It is defined by a single salient loading from Equation Building. Both Quantitative and Number Series are nonsalients. This proposed group factor is able to explain 2.2 percent of the total variance. Overall, the four group factors at the oldest age level are able to explain 14.1 percent of the total variance.

Overall, several observations can be made from the review and technical data. The presence of a strong 'g' component at all three age levels is undeniable. This supports the view that the SBIV is a good measure of one's global ability and that the Composite Score is a useful index (Osberg, 1986; Sandoval and Irvin, 1988; Slate, 1986 and Vernon, 1987).

It is also evident that the factor structure is different for the various ages. Thorndike, Hagen and Sattler (1986b) identify; two group factors at the 2 to 6 years age level (i.e., Verbal and Abstract/visual), three at the 7 to 11 years age (i.e., Verbal, Memory, and Abstract/visual), and four at the 12 to 23 years age level (Verbal, Abstract/visual, Memory and Quantitative). Importantly, the third component of the middle group and the fourth component of the oldest group are singlets. Therefore the four cognitive areas of ability are not found at any age level in the standardizations sample data.

The group factors that are defined by more than one salient loading are typically small. The largest group factor is able to account for nine percent of the total variance. However, this is the exception and most account for less than five percent. Overall, the group factors do not contribute to more than 14.1 percent of the total variance at any age level.

Vernon (1987) correctly observed that many of the tests in the standardization data demonstrated poor relationships with other tests in their hypothesized Reasoning Areas and failed to load on their designated factors. For example Bead Memory is not observed to load along with other memory variables. It is, however, observed to be a salient along with Copying at the youngest age level. Absurdities failed to reach a salient level in both younger groups. Copying and Matrices are nonsalients on the Abstract/visual component identified in the 7 to 11 years group. Matrices is also a nonsalient on this factor in the oldest group. Both Quantitative and Equation Building are nonsalients on the Quantitative Factor that is identified in the oldest subsample. Other violations of hypothesized relationships are present.

A final criticism of the technical data is that it does not evaluate intermediate levels in the hierarchy. No evidence is presented to support or refute the Crystallized or Fluid ability distinction that is made at the second level. Rather, only the first and third levels of the hierarchical model are evaluated.

Osberg (1986), Sandoval and Irvin (1988), Slate (1986), and Vernon (1987) concluded, therefore, that although there is adequate validity for the Composite Score, the hierarchical theory is inadequately confirmed. Slate (1986) states that it makes little sense to him why the test publishers continue to market the SBIV based on a theory "when the data suggest only partial support, at best" (p. 3).

Osberg (1986) notes that now, even the SBIV publishers appear to be suggesting that the hierarchical model was proposed simply to guide test development (Sandoval and Irvin, 1988). Thorndike, Hagen and Sattler (1986b), however, receive some support from the *Standards* for this position. The *Standards* stress that the "construct of interest for a particular test should be embedded in a conceptual framework, no matter how imperfect that framework may be" (American Psychological Association, 1985, pp. 9-10). The test authors are also correct in proposing a detailed theoretical framework. According to the *Standards* the theoretical framework should specify the meaning of a construct, what sets it apart from other constructs and how measures of the construct should relate to other variables. Thorndike, Hagen and Sattler's (1986b) hierarchical model is an attempt to fulfill this requirement.

Ethical Issues and the SBIV

The *Standards* are also clear that construct validation is typically ongoing. It begins at the test development stage and continues "until the pattern of empirical relationship between test scores and other variables clearly indicates the meaning of test score(s)" (American Psychological Association, 1985, pp. 9-10). Therefore, construct validation is viewed as a process and not usually confined to a single study/result.

Despite this recognition that construct validation is not static Thorndike, Hagen and Sattler (1986b) are obliged to highlight the limitations of their existing validity data. They are also required to provide validity data for the major types of theoretical inferences (e.g., Crystallized vs. Fluid abilities), and if these areas have not been explored or empirically established it should be made clear. In addition, Thorndike, Hagen and Sattler (1986b) are compelled to demonstrate that single test scores measure the theoretical constructs they are purported to represent. Both test manuals and promotional material are required to guard against potential misuses.

The earlier review of the technical data suggests that such cautions are warranted. Although Composite Score interpretations are supported, the evidence for the Fluid and Crystallized ability distinction is lacking. Evidence for the four cognitive areas of ability is tenuous and seen to vary between age groups. Specific subtests are also observed to load on factors outside their designated area and fail to reach salient levels in assigned areas (e.g., Bead Memory, and Absurdities). These observations and the fact that test results may have significant implications for examinees provides the rationale for subsequent factor analytic work and this study.

Empirical Reviews and the SBIV

Fritzke (1988) examined the SBIV correlational and factorial results of 168 children referred to an Edmonton clinic. Fritzke's (1988) exploratory factor analytic results are supportive of a strong 'g' component. At the youngest age (6-0 to 8-11 years) level analyzed a second component is supported by both eigenvalue and scree criteria. This is identified as a Verbal ability under quartimax rotation and a Perceptual organization factor under varimax rotation. Importantly, Bead Memory is identified as falling under this latter factor. Fritzke also identifies a three component varimax solution that appears to measure Fluid intelligence, Verbal ability and the ability to Perceive and organize material.

Fritzke's quartimax rotations for his middle group (9-0 to 12-11 years) are supportive of a strong 'g' factor. A second factor, largely Abstract/visual, is partially supported. However, Bead Memory and Memory for Objects are observed to be related to this factor. Upon rotation, Fluid and Crystallized abilities emerged as did a Short-term memory factor.

Fritzke's third and oldest (13 years +) group results are extremely heterogenous. Scree and eigenvalue criteria are supportive of a single, general component. This results is observed to hold under various rotations.

Overall, Fritzke's results are in many ways similar to the SBIV technical data. Support is found for a strong general component. The four cognitive areas of ability are not consistently identified. Finally, specific tests are observed to load outside their hypothesized area.

Sattler (1988) conducted a varimax analysis on the original SBIV standardization data. He identifies two factors at the 2 to 7 year level. These are Verbal Comprehension and Nonverbal Reasoning/Visualization. A third factor is identified as Memory. Several subtests are observed to load outside their cognitive areas (e.g., Memory for Sentences, Quantitative and Bead Memory).

Keith et al (1988) have employed confirmatory factor analytic methods on the standardization sample data. Both the SBIV model and a less restrictive model that remains true to test percepts are evaluated. The latter model allowed for Memory for Sentences, Bead Memory and Absurdities to load on the Abstract/visual factor. Four age groups are analyzed (3-6, 7-11, 12-23 and 3-23 years).

The results for the youngest groups are constrained by the fact that only one test (Quantitative) forms the basis of the Quantitative Reasoning Area Score. Despite adjusting for this problem Keith et al (1988) had considerable difficulty approximating the SBIV model. They proposed a non-memory model and found an "excellent fit to the 2 through 6 year old standardization data. The correlations between the Verbal and Quantitative factors (.63) are somewhat lower than the Verbal and Abstract/visual correlations (.75) and the Quantitative and Abstract/visual correlations (.80).

The confirmatory factor analysis results for 7 to 11 year old age group indicate that, in general, the strict theoretical model fits the standardization data. However, a perfect correlation between the Quantitative and Abstract/visual factors is found possibly suggesting these factors are measuring the same thing. Similar to the younger group the relaxed model is observed to be a better fit. A reduction in the perfect correlation between Quantitative and Abstract/visual is found. The obtained correlation (.97) is still higher than the Verbal-Quantitative correlation (.82).

Similar to the two youngest groups the relaxed model is a better fit for the 12 to 23 age group. When specific subtests are allowed to load on more than one factor," Bead Memory loaded more highly (.52) on the Abstract/visual than on the Memory (.25) factor. Absurdities loaded almost equally on the Verbal (.37) and the Abstract/visual (.40) factors, and Memory for Sentences loaded only slightly higher on Verbal (.44) than on the Memory (.38) factor. As in the strict Binet model, the correlation between Quantitative and the Abstract/visual factor was the highest (.92); the Quantitative-Verbal correlation was somewhat lower (.83)" (p. 265-266).

The confirmatory results for the entire standardization sample are consistent with the findings from the various age groups. A strong 'g' component is found along with the four cognitive areas of ability in the restricted model. Again, however, the "correlation between verbal and quantitative was the lowest factor intercorrelation (.79) ... and the quantitative factor correlated most highly (.92) with the abstract/visual factor ..." (Keith et al., 1988, p. 261).

The results from the relaxed model are also similar to the various age group findings. A significantly better model fit is found when Memory for Sentences, Bead Memory and Absurdities are allowed to load outside their hypothesized Reasoning Areas. Bead Memory is observed to lad highest on the Abstract/visual component. Memory for Sentences for salient loadings on both the Verbal and Short-term memory components. Absurdities loaded highest on the Verbal factor but reached a near salient level on the

Abstract/visual component. Again, the correlation between Abstract/visual and Quantitative factors is high.

Keith et al (1988) conclude that, in general, their results are supportive of the four areas of cognitive ability. The one exception is the lack of evidence for a separate Memory factor at the youngest age. The presence of 'g' is supported. Correlations between the various Area Score are, however, less supportive of the Crystallized-Fluid abilities distinctions. Keith et al (1988) suggest that caution should be used in interpreting SBIV data at this level.

The objective of this study was to further evaluate the proposed SBIV empirical and factor structure using data collected outside the standardization sample. The four cognitive areas of ability and hierarchical arrangement were of special interest. An ancillary purpose of this study was to determine if the SBIV model held across age groups. If the SBIV model was not appropriate alternatives were to be offered. Specific suggestions for test interpretation were to be provided.

A clinic sample of 371 individuals ranging in age from 3 to 23 years was utilized. All ability levels were included and a sizable proportion of the sample consisted of visible minorities. Testing was conducted by graduate students in educational psychology under the supervision of senior psychologists.

The obtained descriptive statistics were remarkably consistent to those proposed by Thorndike, Hagen and Sattler (1986b). Individual test means and standard deviations typically fell within two points of their hypothesized values. Area mean and standard deviation scores were slightly higher but all fell within 5 points of their values. Overall, the clinic sample typically scored approximately 2 points higher and ranged 3 points more than the reported SBIV values.

Data was separately analyzed for each of Thorndike, Hagen and Sattler's age groups. The obtained descriptive statistics were generally similar for all three age groups.

The descriptive data was supportive of the Clinic sample and its use in subsequent analyses.

An early finding was that the age based and optional test selection format for noncompulsory tests must be considered when using the SBIV for research and/or comparative purposes. The possibility of separate analyses using differing combinations of tests was evident. It was recommended that steps be routinely taken to ensure that all available tests are completed and that populations are similar both in age and ability levels prior to making statements of equivalencies. Multiple regression techniques were viewed as the procedure of choice to predict missing scores. However, several subtests were dropped from this analysis given the low number of subjects that actually were administered them.

The problem of missing test data was viewed as being more central to factor analytic studies. Such analyses are based on the intercorrelations between tests. Extreme correlation values may occur when specific pairs of tests are given to selected sample members. Therefore, caution must be used in interpreting SBIV factor analytic studies that have failed to address this issue.

The first factor analytic finding is the presence of a strong General factor in all three subsamples. The unrotated principal component loadings indicate this factor accounts for 47.3 percent of the total variance in the 3 to 6 years group, 58.5 percent of the total variance in the 7 to 11 years group and 72.8 percent of the total variance in 12 to 23 years group. The presence of a strong 'g' is consistent with the SBIV theoretical model and standardization sample data (Thorndike, Hagen and Sattler, 1986b). Additional support for this component is found in factor analytic work on earlier (Burt and John, 1942(a)(b); Hallahan, Ball and Payne, 1973) and current (Fritzke, 1988; Keith et al., 1988) editions of the Stanford-Binet.

The second major finding is that the factor analytic results appear not only to change under rotational method and number of factors extracted but also with age. Findings from several areas in the study including the unrotated principal components indicate that the strength of 'g' increases directly with age. The orthogonal Procrustes randomization test results also suggest developmental influences may be at work. Table 5.0 displays alternative developmental models for each of the three age groups. Where multiple factor loadings exist, tests are aligned according to their highest salient factor loading. For illustrative purposes the presence of 'g' is also highlighted.

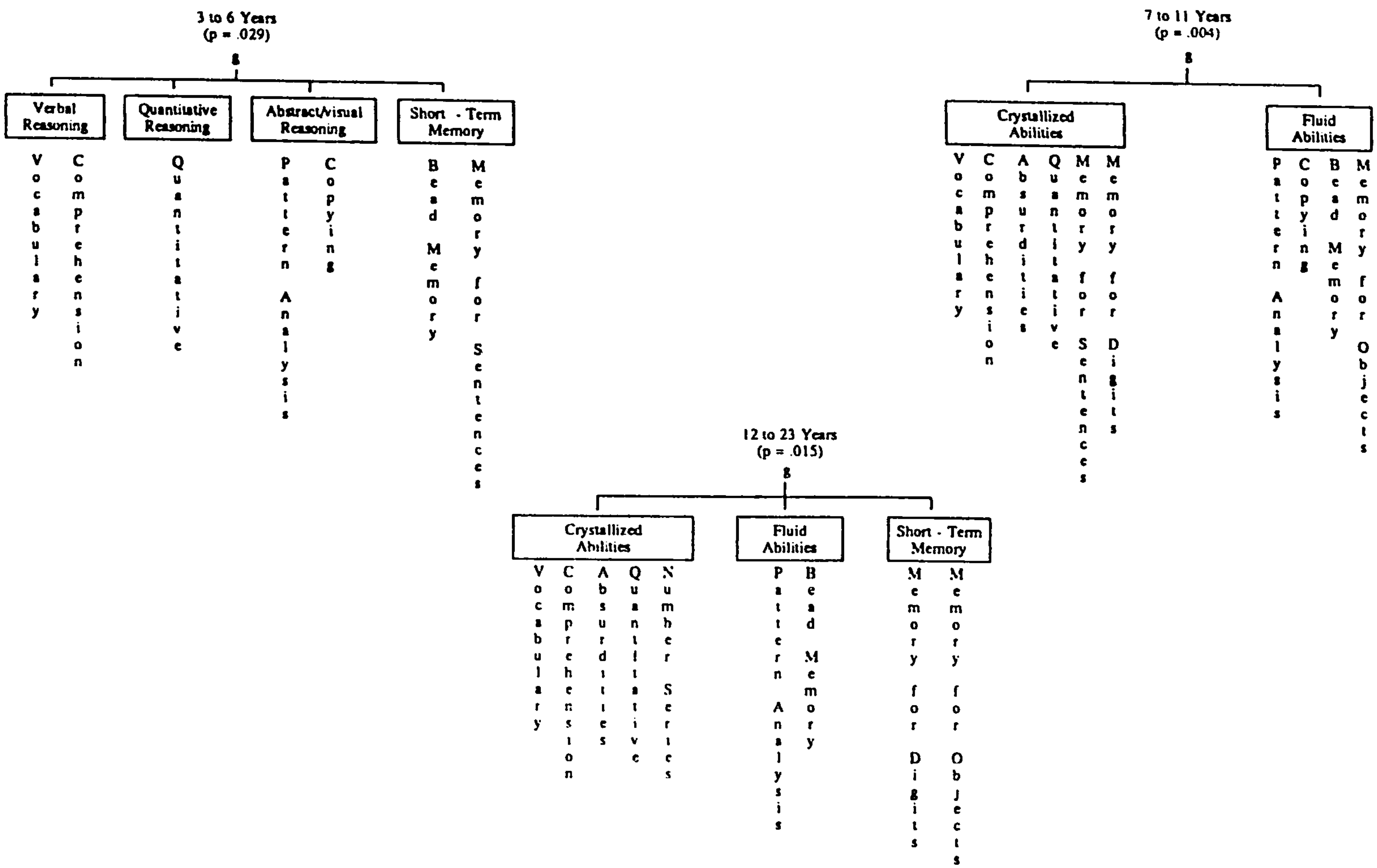
All three developmental models are observed to adhere to the SBIV theoretical structure. Under this analysis, however, the 3 to 6 years group is best described by the four cognitive areas of ability (p=.029). The Quantitative Reasoning Area is the most tentative simply because a single Area test is available for this age group. Despite the fact that most variables load on more than one factor all tests have their highest loading in their designated Areas.

For the youngest group, Comprehension has its highest loading on the Verbal Reasoning factor (.78) but also loaded on Short-term memory (.39). Pattern Analysis loaded highest on the Abstract/visual Reasoning Area (.65) and loaded somewhat less on the Quantitative factor (.57). Bead Memory loaded strongly on the Short-term memory factor (.87) and moderately on the Quantitative factor (.34). Memory for sentences had its strongest loading on the Short-term memory factor (.62) and loaded only somewhat less on the Verbal factor (.52). The existence of multiple loadings and continuities in subsidiary variable-factor relationships is consistent with Keith et al's (1988) findings.

The developmental model that best fits the 7 to 11 years subsample is based on a Crystallized-Fluid abilities differentiation (p=.004). Similar to the youngest subsamples all variables are observed to adhere to their hypothesized substructure. That is, they have their highest loading on their designated factor. Also, like the youngest subsample developmental model the majority of variables in this age group have multiple salient loadings. Absurdities has its highest loading on the Crystallized factor (.59) but also loads

TABLE 5.0

SBIV Developmental Models



in a salient fashion on the Fluid component (.46). Similarly, Quantitative loads highest on the Crystallized ability (.68) and has a lower salient loading on the Fluid factor (.39). Pattern Analysis has salient loadings on both the Fluid (.69) and Crystallized (.50) components. The memory variables are also aligned in a theoretically consistent fashion. Bead Memory and Memory for Objects have their highest loadings on the Fluid (.85 and .77 respectively) as opposed to the Crystallized (.61 and .58) abilities. Whereas the two remaining memory variables have single salient loadings on the Crystallized component. This model also received some support from the two factor multiple group solution and the hierarchical analysis.

A three factor developmental model is best able to describe the oldest subsample (p=.015). Crystallized, Fluid and Short-term memory factors identified. Unlike the two younger subsamples several variables have their highest salient loading outside their theoretical areas. Matrices is observed to have its highest loading on the Crystallized component (.67) and lower loadings on the Fluid (.44) and Short-term memory (.43) components. Bead Memory has its highest loading on the Fluid component (.56) and lower salient loadings on Short-term memory (.53) and Crystallized abilities (.46). Memory for Sentences has its highest loading on the Crystallized factor (.75) and a lower salient loading on the Short-term memory component (.53). Similar to the middle subsample the majority of variables have salient loadings on more than one factor. Five of the 11 variables analyzed have salient loadings on all three components. Some support for this model is garnished from the three component multiple-group factor analysis.

The four component solution under the orthogonal Procrustes randomization analysis, is the only solution that reaches significance for all three subsamples. This solution is displayed by subsample in Table 5.1. As the four component solution more closely adheres to the SBIV theoretical tenets it is recognized as the "test model".

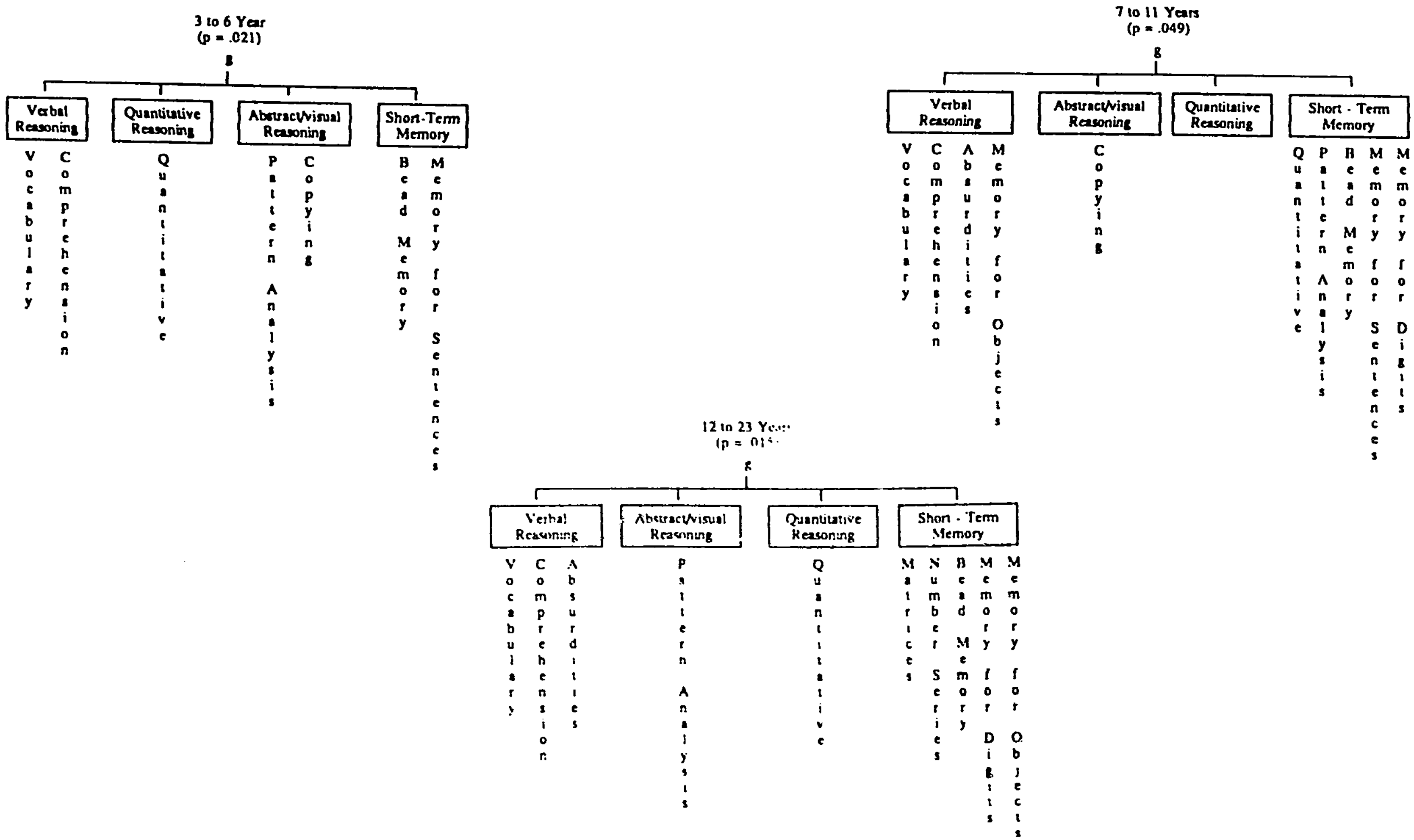
As noted earlier, the four component model is the best overall orthogonal Procrustes solution for the 3 to 6 years subsample. Table 5.1 indicates that all tests at this age have their highest salient loadings on their designated factors. Variables that have salient loadings on additional factors have been outlined earlier. Support for this model for the youngest subsample is also provided by the four component varimax and multiple group solutions. Partial support for this model is provided by the hierarchical results but the presence of a strong 'g' component and the single quantitative test available likely inhibited both Quantitative and Memory factors from emerging in this analysis.

The four component orthogonal Procrustes solution is also significant (p=.05) under the systematic permutations test. However, unlike the youngest subsample and the two factor solution for this age not all tests have their highest loading on their designated factor. This is the third salient finding from this analysis. As can be seen from Table 5.1 the single Quantitative test available has its highest loading on Short-term memory factor (.57). However, this test also has salient loadings on the Quantitative (.37) and Verbal (.36) factor. Memory for Objects has its highest loading on the Verbal factor (.47). This is highly similar to loading on the Abstract/visual (.47) and Short-term memory (.43) components. Pattern Analysis is observed to have its highest loading on the Short-term memory factor (.76). Lower salient loadings are also, however, obtained on the Abstract/visual (.50) and Verbal reasoning (.32) components. Similar to Keith et al's (1988) analysis the presence of multiple salient loadings is striking. Support for this solution is also found in the four component oblique Procrustes, and multiple group solutions. It is also found in the earlier analyses of Fritzke (1988) and Keith et al (1988). Only partial support for this model is, however, found under the exploratory and hierarchical analysis.

The four component solution is a better fit statistically for the oldest group under the orthogonal Procrustes randomization test than most other models with the exception of the three component model. Like the 7 to 11 years subsample several tests have their highest salient loading outside their designated Reasoning Areas. Matrices is observed to have its highest loading on the Short-term memory component (.50). Slightly lower salient

TABLE 5.1

SBIV Test Models



loadings are found on Abstract/visual (.48), Quantitative (.43), and Verbal (.42) components. Similarly, Number Series is observed to have its highest salient loading on Short-term memory (.51). Lower salient loadings are also observed on Verbal (.51), Quantitative (.47), and Abstract/visual (.41) factors. All of the remaining have salient loadings on more than one factor. Additional support for this four component solution is largely obtained from the oblique multiple group analysis. However, only partial support for this model is obtained from most other methods including the exploratory rotations and hierarchical analysis within this study at these ages. The relatively larger influence of *g* for the older group is one possible reason for the limited support for the SBIV four component or hierarchical models. More consistent additional evidence for the four component model for the 12 to 23 years subsample is found in Keith et al's (1988) follow-up analysis of the standardization data.

The varimax, orthogonal Procrustes and multiple group analyses are then, generally supportive of the four cognitive areas of ability across ages. The hierarchical analysis results are perhaps the most disappointing in their ability to offer consistent interpretive advice to SBIV users aside from the presence of an overriding general factor at the highest level. The absence of consistent substructures between these two levels is the fourth major finding of this study.

The provisional acceptance of the 4-1 model has many advantages. Foremost is its basic adherence to two levels of the SBIV theoretical framework. Secondly, this model is loyal to practical SBIV test construction and interpretation designs. Thirdly, this model is able to avoid the complexities of multi-level interpretations that, at this stage, may be ethically risky given their lack of empirical support. A final advantage of this model is its support across all three age groups. Separate models for each age group are cumbersome and particularly problematic when defining age related cut-off points. The provisional acceptance of the SBIV four cognitive areas is restricted by both practical and ethical considerations.

Practical Considerations arising from the SBIV Analyses

Practical considerations are important in choosing to tentatively accept the 4-1 model. Primary interpretative value should continue to be given to the Composite Score. Despite this, the Composite Score should, like any other intelligence measure, never be the sole factor in a program/placement and decision making process. Other factors including clinical history, sensory, capabilities, motivation, emotional status, social variables and alternative opportunities should routinely be considered.

Caution should be utilized in interpretations beyond the Composite Score. This is especially true for the Quantitative Reasoning Area. This cognitive area is the most tenuous of all Area Scores. This is also true for Short term memory at younger ages. Proper qualifications relating to the debate on the existence of these factors should be specified.

Less interpretative weight should be given to Area Scores that are defined by a single subtest. This is particularly problematic for the Quantitative Reasoning Area. At the 3 to 6 years level only one quantitative subtest is available.

Care should be exercised in utilizing Area Scores that are based on subtests known to load outside their designated area. This is especially problematic when it is the only subtest defining that area. When two subtests define an Area Score consideration should be given to whether or not one of these loads outside the Reasoning Area. If this is true the Area Score should be treated as if it is defined by a single subtest. Appropriate interpretative cautions are warranted.

All references to specific subtests should be treated as hypotheses unless empirical evidence exists to substantiate any proposed relationships. Test users should be familiar with the specificity of each subtest. Hypotheses based upon subtest analysis should consider both significant and absolute differences. Test users should also be sensitive to the problems associated with multiple comparisons and the amount of subtest/area variability that can typically be expected (Sattler, 1988). Profile interpretations are further complicated by non-uniform scaled score ranges and test users must be aware of the interpretative limitations of specific scales (Sattler, 1988).

Test users are required to be sensitive to factors influencing SBIV results. These include socioeconomic and educational opportunities. Linguistic and cultural circumstances are often influencing variables that should be considered as subject to change. Test users should be familiar with available SBIV research with special populations. A critical reading of the SBIV 'Technical Manual' is essential to understanding the tests strengths and weaknesses. Supplementary review research is also required to gain a complete understanding of the SBIV.

Ethical Considerations arising from the SBIV Analyses

The qualified acceptance of the SBIV and its four areas of cognitive ability has important and meaningful ethical implications. Ethical standards relevant to the SBIV construct validation process are provided in Chapter 2 (Table 2.1). These are directly related to the previously outlined practical considerations. In keeping with these standards inferences can only be made from the Composite Score (Standard 1.1). More restricted inferences can only be made from the Verbal and Abstract Reasoning Scores but only when relationship have been empirically demonstrated (Standards 7.1 and 8.11). The evidence for a Short term memory area is stronger at two older levels. Propositions from the Quantitative Reasoning Area are likely the most tenuous ethically and this is particularly true at the 3 to 6 years level.

Only the most restricted inferences can be ethically provided from subtest scores. Subtest scores that routinely load on nondesignated Reasoning Areas should be interpreted with appropriate coveats (Standard 1.10) and with respect to their empirical as opposed to theoretical framework. The interpretation of subscores, score differences, or profiles should be extremely limited (Standard 1.3 and 1.9). The evidence for such SBIV interpretations at present is lacking.

Ethical standards relevant to the test manual and promotional materials are central. Presently interpretation of subtest scores and area scores may routinely occur without empirical justification. SBIV test manuals should explicitly state the limitations of both specific Area and subtest scores. The test manual should clearly provide warnings that would guide against misinterpretations. Revisions in both the SBIV test manual and protocols should occur to reduce the possibility of unsubstantiated interpretations (Standard 5.7). When new research (e.g., Sattler, 1988) specifically contradicts or restricts proposed interpretations test users should be alerted through publication houses as well as scholarly journals. It is essential to highlight that while test authors have a strict obligation to adhere to ethical standards this does not remove the primary responsibility from clinicians to ensure all inferences are empirically substantiated. All SBIV test findings that are to be used in the setting educational objectives and instructional strategies must have appropriate empirical referents (Standard 8.11). In summation demands on both test developers and users for diagnostic clarity and specificity are increasing. This does not, however, obviate ethical considerations and needs of the individual which Binet, himself placed first.

Intelligence tests including the SBIV should, therefore, continue to be regarded as fallible instruments. One issue that unites intelligence test researchers, practitioners and consumers is the need for ongoing research (Burt , 1977). This is especially true when an instrument is new or undergoes substantial revision. New or revised measures of intelligence such as the SBIV are catalysts for exploration and re-evaluation both within the narrow confines of intellectual specification and the broader parameters of theoretical clarification. A substantial research effort is required even within the narrow parameters of SBIV construct validity.

Future Research Considerations

Contemporary intelligence researchers are also as much concerned about what intelligence tests don't measure as they are about what they do measure. Binet's early position is in line with this view. He offered his measure simply as a means of classifying students along a continuum of need for special education programming.

A second area where modern research psychologist are aligning with Binet is that intelligence is only one aspect of an individual's personality and, as such, it "cannot be assessed and described in isolation" (Matarazzo, 1972, p. 20). Trends in the search for multiple intelligences are an extension of this view. So too is the growing view that conceptions of intelligence are often affected by social and political contingencies.

Progress in understanding intelligence is, therefore, being achieved in research that transcends traditional boundaries of theory, field of inquiry and psychometrics (Estes, 1986). Sternberg and Weinberg (1986) stress that much of the early psychological research is dominated by the study of intelligence tests" (p. ix). Recently, there is growing evidence of a resurgence of theoretical issues that question the very nature of the constructs upon which these instruments are based. Much research will likely continue to debate the correct number, arrangement and inter-workings of fundamental human abilities (Steinberg, 1977). Attention will likely focus on those theories that are able to offer an rapprochement between diverse findings (e.g., hierarchical theories).

One traditional form of demarcation between intelligence theorists is the focus on individual differences as opposed to more contemporary theorists who focus on task differences. The latter emphasize information processing and learning strategies. Measurement of these processes has been greatly enhanced by high speed computers and other advanced technologies. Future research is likely to involve greater recognition and integration of both schools of thought.

A third area of research that is gaining momentum is the focus on adaptive intelligencies and environmental influences (Gardner, 1983 and Sternberg, 1985). Such theorists emphasize the contextual aspect of "intelligent" behavior and sociological variables. Emphasis is often placed upon general, real-life problem-solving abilities rather than relatively more narrow "artificial" measures.

A second aspect of the socio-cultural perspective is the issue of heritability. Gould (1981) stresses that there are two primary falsehoods associated with hereditarianism. First, is the equation of heritable with inevitable. Gould states that much research is required to identify the limits of human malleability. The second falsehood is the emphasis by hereditarians on between group as opposed to within group differences. The latter are generally perceived to be much larger than the former, a finding Gould (1981) stresses that is often ignored by those who advocate racial intelligence theories. Ancillary to this work is the relatively unexplored areas of age-linked competencies and developmental transformations (Anastasi, 1988b).

A fourth and relative area of research focuses on behavioral genetics and/or the neuropsychological processes involved in intelligence. Presently, research is growing in the area of genetics as attempts are being made to map the entire human genetic structure. Efforts are also being made to isolate specific genes and determine their function. The ultimate goal of such research is to explore the interaction between genetics and human behavior in order to promote human development in all spheres.

Advances in neuroendocrinology and clinical electrophysiology (e.g., positron emission tomography (P.E.T.) and magnetic resonance imaging (M.R.I.) are promising for the study of major cognitive and affective processes. Levin (1987) states that P.E.T. scans and magnetic resonance spectroscopy may soon become dependent variables in the evaluation of regional cerebral metabolism and activation procedures. Intelligence research is likely to be greatly enhanced as the neurological bases for specific abilities, information processes, decision-making, memory and attention/concentration are delineated (Levin, 1987).

The future of factor analysis is also bright. Confirmatory factor analysis and causal modeling are likely to make significant contributions to intelligence research. Additional advances are likely to be made in the areas of factor replication, invariance and robustness (Gorsuch, 1983). Gorsuch (1983) emphasizes that "as knowledge of principles for designing good factor analytic studies spreads, factor analytic research can be expected to contribute substantially to psychology and other similar fields" (p. 336).

The construct of human intelligence will likely continue to occupy a central role in psychological theory, research and practice. The SBIV and its predecessors have contributed greatly to all of these areas. "New theories and vigorous research agendas hold promise for an increased understanding of the nature of intellectual development across the life span, more effective tools for assessing individual differences in intellectual competence and an expanded awareness of how society can intervene to enhance the individual's intellectual skills (Weinberg, 1989, p. 103).

It is clearly evident then, that contemporary research views are remarkably similar to those of Alfred Binet. Therefore, one can easily agree with Sternberg (1986) that a good first step in future research is a reading of early theorists. A careful reading of Alfred Binet's work is an excellent beginning. The SBIV is a logical outgrowth of these early views and is a culmination of over eighty years of theoretical and factorial pursuit. Similar to its predecessors it is likely to stimulate considerable research and contribute greatly to both the theoretical and empirical understanding of human intelligence.

REFERENCES

American Psychological Association. (1985). Standards for educational and psychological testing. Washington, DC: American Psychological Association, Inc.

Anastasi, A. (1982). Psychological Testing, (5th ed.). New York: MacMillan Publishing Co., Inc.

Anastasia, A. (1988a). Psychological Testing, 6th ed. New York: Macmillan Publishing co., Inc.

Anastasi, A. (1988b). Review of the Stanford Binet Intelligence Scale, Fourth Edition. In Conoley, J. C.; Kramer, J. J., & J. V. Mitchell Eds. (1988). The Supplement to the Ninth Mental Measurements Yearbook. Lincoln, NE: University of Nebraska Press, 198-200.

Baggaley, A.R. (1982). Deciding on the ratio of number of subjects to number of variables in factor analysis. Multivariate Experimental Clinical Research, 6, 81-85.

Baird, D. (1987). Exploratory factor analysis, instruments and the logic of discovery. British Journal for Philosopy of Science, 38, 319-337.

Barnard, J. (1987). Consideration regarding the use of the new Stanford-Binet. The Alberta School Psychologist, 7(1), 16.

Binet, A., & Henri, V. (1905). La psychologie individuelle. L'Annee Psychol., 2, 411-465. cited in Frank, G. (1983). The Wechsler enterprise: An assessment of the development, structure, and use of the Wechsler tests of intelligence. New York: Pregamon Press.

Binet, A., & Simon, T. (1905a). Upon the necessity of establishing a scientific diagnoses of inferior states of intelligence. In E.S. Kite, 1916 (Trans.), The development of intelligence in children (pp. 9-36). New Jersey: Vineland.

Binet, A., & Simon, T. (1905b). New methods for the diagnosis of the intellectual level of subnormals. In E. S. Kite, 1916 (Trans.), The development of intelligence in children (pp. 37-90). New Jersey: Vineland.

Binet, A., & Simon, T. (1908). The development of intelligence in the child. In E.S. Kite, 1916 (Trans.), The development of intelligence in children (pp. 182-273). New Jersey: Vineland.

Binet, A., & Simon, T. (1911). New investigations upon the measure of the intellectual level among school children. In E.S. Kite (Trans.), 1916, The development of intelligence in children (pp. 274-329). New Jersey: Vineland.

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

Broadbooks, W.J., & Elmore, P.B. (1987). A Monte Carlo study of the sampling distribution of the congruence coefficient. Educational and Psychological Measurement, 47, 1-11.

Brody, E.B., & Brody, N. (1976). Intelligence: Nature, determinants, and consequences. New York: Academic Press.

Burt, C. (1939a). The Relations of Educational Abilities. British Journal of Educational Psychology, IX, 45-71.

Burt, C. (1939b). The Latest Revision of the Binet Intelligence Tests. Eugenics Review XXX, 255-260.

Burt, C. (1940). The factors of the mind. London: University of London Press.

Burt, C. (1948). Factor analysis and canonical correlation. British Journal of Psychology, Statistical Section, 1, 95-106.

Burt, C. (1949). The distribution of intelligence. British Journal of Educational Psychology, 19, 190.

Burt, C. (1955). The evidence for the concept of intelligence. In H.J. Eysenck, (Ed.), 1973, The measurement of intelligence (pp. 3-22). Lancaster, England: Medical and Technical Publishing Co.

Burt, C. (1977). The subnormal mind (3rd ed.). London, England: University of London Press.

Burt, C., & John, E. (1942a). A factorial analysis of the Terman Binet tests: Part I. British Journal of Educational Psychology, 12, 117-127.

Burt, C., & John, E. (1942b). A factorial analysis of the Terman Binet tests: Part II. British Journal of Educational Psychology, 12, 156-161.

Burt, C.L. (1917). Distribution and relations of educational abilities. London: P.S. King.

Butcher, H.J. (1968). Readings in human intelligence: Its nature and assessment. London: Metheun.

Butterfield, E.C. (1986). Intelligent action, learning and cognitive development might all be explained with the same theory. In R.J. Sternberg & D..K. Detterman (1986), (Eds.) What is intelligence? Contemporary viewpoints on its nature and definition. New Jersey: Norwood.

Carajaval, H., Hardy, K., Harmon, K., Sellers, T., et al. (1987b). Relationships among scores on the Stanford-Binet IV, Peabody Picture Vocabulary Test-Revised, and Columbia Mental Maturity Scale. Bulletin of the Psychomonic Society, 25(4), 275-276.

Carajaval, H., McVey, S., Sellers, T., Weyand, K., et al. (1987a), Relationships between scores on the General Purpose Abbreviated Battery. Columbia Mental Maturity Scale and goodenough-Harris Drawing Test. Psychological Record, 37(1), 127-130.

Carroll, J.B. (1985). Exploratory factor analysis: A tutorial. In D.K. Detterman (Ed.), Current topics in human intelligence: Research methodology. (vol. 1) (25-28). Norwood, NJ: Ablex Publishing Co.

Carvajal, H. & Weyand, K. (1986). Relationships between scores on ths Stanford-Binet IV and Wechsler Intelligence Scale for Children-Revised. Psychological Reports, 59, 963-966.

Cattell, R.B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. Journal of Educational Psychology, 54, 1-22.

Cattell, R.B. (1971). Abilities: Their structure, growth, and action. Boston: Houghton Mifflin Company.

Cattell, R.B. (1987). Intelligence: Its structural growth and action. New York: Elsevier.

Catte[illegible] Horn, J.L. (1978). A check on the theory of fluid and crystallized i[illegible] with description of new subtest designs. Journal of Educational M[illegible]ment, 15, 139-164.

Clark, L. (1988). The Stanford-Binet Intelligence Scale: Fourth Edition: A Validity Study. Unpublished Masters Thesis: University of Alberta, Edmonton, Alberta.

Cronbach, L. J. (1988). Review of the Stanford-Binet Intelligence Scale, Fourth Edition in Conoley, J. C., Kramer, J. J., & J. V. Mitchell, Eds. (1988). The Supplement to the Ninth Mental Measurements Yearbook. Lincoln, NE: University of Nebraska Press, 200-203.

Davis, E.E., & Rowland, T. (1974). A replacement for the venerable Stanford-Binet? Journal of Clinical Psychology, 30, 517-521.

Dillon, W.R., & Goldstein, M. (1985). Multivariate Analysis: Methods and Applications. New York: Wyley.

DuBois, P. (1970). A history of psychological testing. Boston Allyn & Bacon.

Dunn, L. M. (1980). Preface to the 1980 reprint edition of Binet, A., & Simon, T. (1980). The Development of Intelligence in Children. Nashville, TN: Williams Printing Company, v-viii.

Edgington, E.S. (1987). (2nd ed.). Randomization tests. New York: M. Dekker Pub.

Estes, W.K. (1986). Memory Storage and retrieval processes in category learning. Journal of Experimental Psychology: General, 115(2), 155-174.

Eysenck, H.J. (1979). The structure and measurement of intelligence. Berlin: Springer-Verlag.

Fagan, T. (1986a). Stanford-Binet draws board attention. Communique, 15(1), 1-2.

Fagan, T. (1986b). Assembly withdraws concern over Binet. Communique, 15(3), 1.

Fancher, R.E. (1985a). Spearman's original computation of g. A model for Burt? British Journal of Psychology, 76, 341-352.

Fancher, R.E. (1985b). The intelligence men: Makers of the IQ controversy. New York: W.W. Norton & Company.

Frank. G. (1983). The Wechsler enterprise: An assessment of the development, structure, and use of the Wechsler tests of intelligence. New York: Pergamon Press.

Freeman, F.S. (1955). Theory and practice of psychological testing (revised edition). New York: Henry Holt and Company.

French, L.A. (1986). MR testing and evaluation: New dimensions and old concerns. Psychology in the Schools, 23, 64-76.

Fritzke, (1988). Validation of the Fourth Edition of the Stanford-Binet IV and the WISC-R. University of Alberta Unpublished Masters Thesis.

Galton, F. (1883). Inquiries into human faculty and its development. London: MacMillan.

Gardner, H. (1983). Frames of mind: The theory of multiple intelligences. New York: Basic Books.

Gardner, H. (1985). Frames of mind. New York: Basic Books Inc.

Garfinkel, R., & Thorndike, R.L. (1976). Binet item difficulty then and now. Child Development, 47, 959-965.

Gorsuch, R.L. (1983). Factor analysis (2nd ed.). Philadelphia: W.B. Saunders.

Gould, S.J. (1981). The mismeasure of man. New York: W.W. Norton and Company.

Grossman, F.M. (1983). Interpreting WISC-R Verbal-Performance discrepancies: A note for practitioners. Perceptual and Motor Skills, 56, 96-98.

Guilford, J.P. (1936). Psychometric methods. New York: McGraw Hill Book Co.

Guilford, J.P. (1954). Psychometric methods (2nd ed.). New York: McGraw Hill.

Guilford, J.P. (1956). Fundamental statistics in psychology and education. New York: McGraw Hill.

Guilford, J.P. (1967). The nature of human intelligence. New York: McGraw Hill.

Guilford, J.P. (1977). The invariance problem in Factor analysis. Educational and Psychological Measurement, 37, 11-19.

Guilford, J.P. (1977). The invariance problem in factor analysis. Educational and Psychological Measurement, 37, 11-19.

Guilford, J.P. (1981). Factor analysis. Educational and Psychological Measurement, 37, 11-19.

Guilford, J.P. (1981). Higher-order structure-of-intellect abilities. Multivariate Behavioral Research, 16, 411-435.

Guilford, J.P. (1982). Cognitive psychology's ambiguities: Some suggested remedies. Psychological Review, 89, 48-59.

Guilford, J.P. (1985). The structure-of-intellect model. In B.B. Wolman (Ed.), Handbook of intelligence: Theories, measurements, and applications (pp. 225-266). New York: John Wiley & Sons.

Guilford, J.P., & Hoepfner (1971). The analysis of intelligence. New York: McGraw Hill.

Gustafsson, J.E. (1984). A unifying model for the structure of intellectual abilities. Intelligence, 8, 179-203.

Guttman, L. (1956). "Best Possible" systematic estimate of communalities. Psychometrika, 21, 273-285.

Hallahan, D.P., Ball, D.W., & Payne, J.S. (1973). Factorial composition of the short form of the Stanford-Binet with culturally disadvantaged Head Start children. Psychological Reports, 32, 1048-1050.

Harman, H.H. (1967). Modern factor analysis. Chicago: The University of Chicago Press.

Harris, A.J. & Shakow, D. (1937). The clinical significance of numerical measures of scatter on the Stanford-Binet. Psychological Bulletin, 34, 134-150.

Heeler, R.M., & Whipple, T.W. (1976). A Monte Carlo aid to the evaluation of maximum likelihood factor analysis solutions. British Journal of Mathematical and Statistical Psychology, 29, 253-256.

Hendrickson, A.E. & White, P.O. (1966). A method for rotation of higher-order factors. British Journal of Mathematical Statistical Psychology, 19, 97-103.

Hill, T.D.,Reddon, J.R., & Jackson, D.N. (1985). The factor structure of the Wechsler scales: A brief review. Clinical Psychology Review, 5, 287-306.

Hofman, R. (1988). Comparability of Binet (4th) Full Scale and Abbreviated IQs. Communique, 17, 27.

Hofman, R. (1988). Comparability of Binet (4th) full scale and abbreviated IQs. Communique, (May), 27.

Holden, R.R., Reddon, J.R., Jackson, D.N. & Helmes, E. (1983). The construct heuristic applied to measurement of psychopathology. Multivariate Behavioral Research, 18(1), 37-46.

Holroyd, R.G., & Bickley, J. (1976). Comparison of the 1960 and 1972 revisions of the Stanford-Binet L.M. Journal of Youth and Adolescence, 5, 101-104.

Horn, J.L. (1967). On subjectivity in factor analysis. Educational and Psychological Measurement, 27(4), 811-820.

Horn, J.L., & Cattell, R.B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. Journal of Educational Psychology, 57, 253-270.

Horn, J.L., & Cattell, R.B. (1982). Whimsy and misunderstandings of Gf-Gc theory: A comment on Guilford. Psychological Bulletin, 91, 623-633.

Humphreys, L.G. (1962). The organization of human abilities. American Psychologist, 17, 475-483.

Humphreys, L.G. (1967). A critique of Cattell's theory of fluid and crystallized intelligence: A critical experiment. Journal of Educational Psychology, 58, 120-136.

Humphreys, L.G., Ilgen, D., McGrath, D. & Montanelli, R. (1969). Capitalization on chance in rotation of factors. Educational and Psychological Measurement, 29, 259-271.

Hurley, J., & Cattell, R.B. (1962). The procrustes program: Producing direct rotation to test a hypothesized factor structure. Behavioral Science, 7(2), 258-262.

Jackson, D.N., & Morf, M.E. (1974). Testing the null hypothesis for rotation to a target. Multivariate Behavioral Research, 9, 303-310.

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

Jones, L.Y. (1954). Primary abilities in the Stanford-Binet, age 13. Journal of Genetic Psychology, 84, 125-147.

Kail, R. V., & Pellegrino, J. W. (1985). Human intelligence perspectives and prospects. New York: Freeman Press.

Kaiser, H.F. (1958). The varimax criterion for analytic rotation in factor analysis. Psychometrika, 23, 187-200.

Kaufman, A.S. (1976). A new approach to the interpretation of test scatter on the WISC-R . Journal of Learning Disabilities, 9, 160-168.

Kaufman, A.S. (1979). WISC-R research: Implications for interpretation. School Psychology Digest, 8, 5-27.

Kaufman, A.S. & Kaufman, N.L. (1983). Kaufman assessment battery for children. Circle Pines MN: American Guidance Service.

Keith, T.Z. (1987). Assessment research: An assessment and recommended interventions. School Psychology Review, 16(3), 276-289..

Keith, T.Z., Cool, V.A., Novak, C.G., White, L.J., & Pottebaum, S.M. (1988). Confirmatory factor analysis of the Stanford-Binet Fourth Edition: Testing the theory-test match. Journal of School Psychology, 26, 253-274.

Kramer, J.J., Henning-Stout, M., Ullman, D.P. & Schellenberg, R.P. (1987). The viability of scatter analysis on the WISC-R and the SBIS: Examining a vestige. Journal of Psychoeducational Assessment, 5, 37-47.

Levin, H.S. (1987). Introduction in Levin, H.S., Grafman, J. and Eisenberg, H.M. (Eds) (1987). Neurobehavioral recovery from head injury. New York: Oxford University Press, pp. 1-10.

Loo, R. (1979). The orthogonal rotation of factors in clinical research: A critical note. Journal of Clinical Psychology, 35, 762-765.

Lorr, M., & Meister, R.K. (1941). The concept of scatter in the light of mental test theory. Education and Psychological Measurement, 1, 303-310.

Lubin, B., Larsen, R.M., Matarazzo, A.D. (1984). Patterns of Psychological test usage in the United States: 1935-1982. American Psychologist, 39, 441-454.

Lukens, J. (1988). Comparison of the fourth Edition and the L-M Edition of the Stanford-Binet Used with Mentally Retarded Persons. Journal of School Psychology, 26(1), 87-90.

Lukens, J. (1988). Comparison of the fourth edition and the L-M edition of the Stanford-Binet used with mentally retarded persons. Journal of School Psychology, 26(1), 87-89.

Maguire, T.O. (1986). Applications of new directions in statistics to educational research. Alberta Journal of Educational Research, 32(2), 154-171.

Matarazzo, J.D. (1972). Wechsler's Measurement and Appraisal of Adult Intelligence. Baltimore: Williams & Wilkens.

Matarazzo, J.D., Daniel, M.H., Prifitera, A. & Herman, D.O. (1988). Inter-subtest scatter in the WAIS-R standardization sample. Journal of Clinical Psychology, 44(6), 940-950.

Mayr, E. (1982). The growth of biological thought. Cambridge: Belknap Press.

McDonald, R.P. (1985). Factor analysis and related methods. Hillsdale, NJ: Lawrence Erlbaum Associations.

McDonald, R.P., & Muliak, S.A. (1979). Determinancy of common factors: A nontechnical review. Psychological Bulletin, 1986, 297-306.

McNemar, Q. (1942). The revision of the Stanford-Binet Scale: An analysis of the standardization data. Boston: Houghton Mifflin Company.

McNemar, Q. (1964). Lost: Our intelligence? Why? American Psychologist, 19, 871-882.

Meloff, R. (1987). The Stanford-Binet Intelligence Scale: Fourth Edition: A Descriptive Study. Unpublished Master's Thesis. University of Alberta, Edmonton, Alberta.

Mulaik, S.A. (1972). The foundations of factor analysis. New York: McGraw Hill.

Nesselroade, J.R. & Bates, P.B. (1970). On a dilemma of comparative factor analysis: A study of Factor matching based on random data. Educational and Psychological Measurement, 30, 935-948.

Nesselroade, J.R. & Bates, P.B., & Labouvie, F.W. (1971). Evaluating factor invariance in oblique space: Baseline data generated from random numbers. Multivariate Behavioral Research, 6, 233-241.

Nunnally, J.C. (1967). Psychometric theory. New York: McGraw Hill.

Nunnally, J.C. (1970). Introduction to psychological measurement. New York: McGraw Hill.

Nunnally, J.C. (1978). Psychometric theory. (2nd ed.). New York: McGraw Hill.

Osberg, D. (1986). New tests: New options. Communique, 15(3), 3.

Paunonen, S.V. (1987). Test contruction and targeted factor solutions derived by multiple group and procrustes methods. Multivariate Behavioral Research, 22, 437-455.

Peterson, J.C. (1925). Early Conceptions and tests of intelligence. New York: World Book Company.

Ramsey, P.H., & Vane, J.R. (1970). A factor analytic study of the Stanford Binet with young children. Journal of School Psychology, 8, 278-283.

Reddon, J.R. (1990). The Rejection of the Hypothesis of Complete Independence Prior to Conducting a Factor Analysis. Multivariate Experimental Clinical Research, 9, 123-129..

Reddon, J.R., Marceau, R. & Holden, R. (1985). A confirmatory evaluation of the Profile of Mood States: Convergent and discriminant item validity. Journal of Psychopathology and Behavioral Assessment, 7(3), 243-259.

Reynolds, C.R. (1988). Putting the individual into aptitude-treatment interaction. Exceptional Children, 54(4), 324-331.

Reynolds, C.R., & Clark, J.R. (1983). Cognitive assessment of the preschool child. In K. Paget and B.Bracken (Eds.), Psychoeducational assessment of the preschool child. New York: Grune and Stratton.

Reynolds, C.R. & Clark, J.R. (1985). Profile analysis of standardized intelligence test performance of very low functioning individuals. Journal of School Psychology, 23, 277-283.

Reynolds, C.R., Kamphaus, R.W. & Rosenthal, B.L. (1988). Factor analysis of the Stanford-Binet Fourth Edition for ages 2 years through 23 years. Measurement and Evaluation in Counseling and Development, 21, 52-63.

Rothlisberg, B.A. (1987). Comparing the Stanford-Binet Fourth Edition to the WISC-R: A concurrent validity study. Journal of School Psychology, 25, 193-196.

Royce, J.R. (1958). The development of factor analysis. Journal of General Psychology, 58, 139-164.

Rummel, R.J. (1970). Applied factor analysis. Evanston: Northwest University Press.

Salvia, J., Ysseldyke, J.E., & Lee, M. (1975). 1972 revision of the Stanford-Binet: A farewell to mental age. Psychology in the Schools, 12, 421-422.

Sandoval, J., & Irvin, M.G. (1988). Review of the Stanford-Binet Intelligence Scale Fourth Edition.

Sandovil, J. & Irvin, M.G. (1988). Stanford-Binet Intelligence Scale: Fourth Edition Technical Manual. CASP Today, 3-5.

Sattler, J.M. (1982). Assessment of children's intelligence and special abilities (2nd ed.). Boston: Allyn and Bacon.

Sattler, J.M. (1988). Assessment of children's intelligence and special abilities (3rd ed.). Boston: Allyn and Bacon.

Sim, J. (1987). The Stanford-Binet IV: Recent criticisms and issues. The Alberta School Psychologist, 7(1), 17-21.

Slate, J. (1986). new does not necessarily mean better. Communique, 15(1), 2-3.

Spearman, C. (1904a). "General Intelligence," objectively determined and measured. American Journal of Psychology, 15, 201-203.

Spearman, C. (1904b). "General Intelligence," objectively determined and measured. American Journal of Psychology, 15, 203-221.

Spearman, C. (1927). The abilities of man: Their nature and measurement. London: MacMillan and Co., Limited.

SPSS. (1988). SPSS-X Users Guide (3rd ed.) Chicago: SPSS Inc.

Sternberg, R.J. (1977). Intelligence, information, processing and analogical reasoning: The componential analysis of human abilities.

Sternberg, R.J. (1979). Intelligence at the interface between differential and cognitive psychology: Prospects and proposals. In R.J. Sternberg, & D.K. Dettermon (Eds.), 1986, Contemporary Viewpoints on Its Nature and Definition (pp. 33-60). New Jersey: Apex.

Sternberg, R.J. (1986). The future of intelligence testing. Educational Measurement: Issues and Practice, 5, 19-22.

Telzrow, C. (1987). New Stanford-Binet and preschoolers. The Alberta School Psychologist, 7(1), 22.

Terman, L.W. (1919). The intelligence of school children, how children differ in ability, the use of mental tests in school grading and the proper education of exceptional children. Boston: Houghton Mifflin.

Terman, L.W. (1960). Stanford Binet Intelligence Scale: Manual for the third revision Form L-M. Boston: Houghton Mifflin.

Terman, L.W., & Merrill, M.A. (1937). Measuring intelligence: A guide to the administration of the new revised Stanford-Binet Test of Intelligence. Cambridge, MA: The Riverside Press.

Terman, L.W. & Merrill, M.A. (1953a). Stanford-Binet Scales. In A.Weiderd (ed.) Contributions toward medical psychology: Theory and psychodiagnostic methods. Vol.1. New York: The Ronald Press Company.

Terman, L.W., & Merrill, M.A. (1953b). 1937 Stanford-Binet Scales (pp. 510-521). In Arthur Weider (Ed.), Contributions toward medical psychology: Theory and psychodiagnostic methods. Vol. II. New York: The Ronald Press Company.

Terman, L.W.., & Merrill, M.A. (1973). Stanford-Binet Intelligence Scale: Manual for the third revision from L-M: 1972 Norms Edition. Boston: Houghton Mifflin Company.

Thompson, J.E. (1984). An empirical versus conceptual model of the Stanford-Binet for interpretive purposes. Unpublished masters degree, University of Alberta, Edmonton.

Thorndike, R.L. (1987). Personal communication to B. Fritzke on validation of the Stanford-Binet Fourth Edition.

Thorndike, R.L., Hagen, E.P. & Sattler, J.M. (1986a). The Stanford-Binet Intelligence Scale: Fourth Edition Guide for Administering and Scoring. San Diego State University, Âiverside Publishing Company.

Thorndike, R.L., Hagen, E.P. & Sattler, J.M. (1986b). The Stanford-Binet Intelligence Scale: Fourth Edition Technical Manual. San Diego State University, Riverside Publishing Company.

Thorndike, R.L., Hagen, E.P. & Sattler, J.M. (1987). The Stanford-Binet Intelligence Scale: Fourth Edition. Chicago: The Riverside Publishing Company.

Thurstone, L.L. (1938). Primary mental abilities. Psychometric Monographs, No. 1. Chicago: University of Chicago Press.

Thurstone, L.L. (1947). Multiple factor analysis: A development and expansion of the vectors of mind. Chicago: University of Chicago Press.

Thurstone, L.L., & Thurstone, T.G. (1941a). Factorial studies of intelligence. Psychometric Monographs, No. 2. Chicago: University of Chicago Press.

Thurstone, T.G. (1941b). Primary mental abilities of children. Educational and Psychological Measurement, *1*, 105-116.

Timm, N.H. (1970). The estimation of variance-covariance and correlation matrices from incomplete data. Psychometrika, 35, 417-437.

Tucker, L.R. (1951). A method of synthesis of factor analysis studies (Personnel Research Section Report No. 984). Washington, DC: Department of the Army.

Tuddenham, R.D. (1962). The nature and measurement of intelligence. In L.J. Postman (Ed.) (1962). Psychology in the making. New York: Knopf, 469-525.

Undheim, J.O. (1976). Ability structure in 10-11-year-old children and the theory of fluid and crystallized intelligence. Journal of Educational Psychology, 4, 411-423.

Undheim, J.O. (1981a). On intelligence I: Broad ability factors in 15-year-old children and Cattell's theory of fluid and crystallized intelligence. Scandinavian Journal of Psychology, 22, 171-179.

Undheim, J.O. (1981b). On Intelligence II: A neo-Spearman model to replace Cattell's theory of fluid and crystallized intelligence. Scandinavian Journal of Psychology, 22, 181-187.

Undheim, J.O. & Horn, J.L. (1977). Critical evaluation of Guilford's structure-of-intellect theory. Intelligence, 1, 65-81.

Velicer, W.F. & Jackson, D.N. (1990). Component analysis versus common factor analysis: some further observations. Multivariate Behavioral Research, 25(1), 97-114.

Velicer, W.F. & Jackson, D.N. (1990). Component analysis versus common factor analysis: some issues in selecting an appropriate procedure. Multivariate Behavioral Research, 25(1), 1-28.

Velicer, W.F., Peacock, A.C., & Jackson, D.N. (1982). A comparison of component and factor patterns: A Monte Carlo approach. Multivariate Behavioral Research, 17, 371-388.

Vernon, P.E. (1953). Current Trends in British Psychology. London: Metheun & Co. Ltd.

Vernon, P.E. (1961). The structure of human abilities. (2nd ed.) London: Methuen & Co. Ltd.

Vernon, P.E. (1979). Intelligence Testing 1928-1978. What next? Edinburgh: Lindsay & Co.

Vernon, P.E. (1987). The demise of the Stanford-Binet Scale. Canadian Psychology, 28, 251-258.

Waddell, D.D. (1980). The Stanford-Binet: An evaluation of the technical data available since the 1972 restandardization. Journal of School Psychology, 18, 203-209.

Walters, J.M., & Gardner, H. (1986). The theory of multiple intelligences: Some issues and answers. In R. J. Sternberg & R.K. Wagner (1986) (Eds.), Practical intelligence: Nature and origins of competence in everyday world. New York: Cambridge University Press.

Weinberg, R.A. (1989). Intelligence and IQ: Landmark Issues and Great Debates. American Psychologist, 44(2), 98-104.

Wolf, T.H. (1969a). The emergence of Binet's conceptions and measurements of intelligence: A history of the creative process. Part 1. Journal of the History of the Behavioral Sciences, 5, 113-134.

Wolf, T.H. (1969b). The emergence of Binet's conceptions and measurements of intelligence: A history of the creative process. Part 2. Journal of the History of the Behavioral Sciences, 5, 207-237.

Wolf, T.H. (1973). Alfred Binet. Chicago: University of Chicago Press.

Wright, R.E. (1939). A factor analysis of the original Stanford-Binet Scale. Psychometrika, 4, 209.220.

Wrigley, & Neuhaus, J.O. (1955). The matching of two sets of factors. American Psychologist, 10, 418.

Yeboah, J. D., & Peat, D. (1986). Towards an integration of psychometric and Vygotskian views of intelligence. Unpublished paper. University of Alberta.

APPENDIX A

Degenerative Hierarchical Solutions

Introduction

The following hierarchical solutions are degenerative. They are degenerative in that results at lower levels preclude meaningful findings at higher levels. Therefore, they are simply outlined to demonstrate their lack of interpretative value.

Degenerative Oblique Hierarchical Solutions for the 3 to 6 Years Subsample.

Tables A.0 to A.13 display the ancillary oblique hierarchical solutions for the 3 to 6 years subsample. The 4-3 solution displayed in tables A.0 to A.2 is unsatisfactory for explanation purposes. All three components are largely General factors. Correlations between all factors are high. However, these are in the negative direction for the second and third components.

The four-three-two solution is displayed in Tables A.3 to A.5 This solution is largely unacceptable given the results at the second level that are outlined above. At the highest level a Verbal-Abstract/visual distinction is apparent. Copying and Memory for Sentences have salient loadings on both components. The moderate negative correlation between the two components is also unexpected given the hypothesized *'g'* linkages between abilities.

The four-three-two-one results are displayed in Table A.6. A higher level Abstract reasoning component is identified. Importantly, all three negative salient loadings that define this component are from separate cognitive Reasoning Areas. Therefore, this component is set apart from the SBIV theoretical underpinnings and is rejected both at the three or well or single component levels.

The four-three-one oblique model is presented in Table A.7. A bipolar component is found. At one end is the Abstract reasoning component identified in the four-three-two-one solution. At the other is a Verbal theme. Again, this solution is largely unacceptable at both the second and highest level.

The four-two oblique solution is presented in Tables A.8 to A.10. Both factors are General and largely indistinguishable. The congruence coefficients are less than acceptable. The large negative correlation between factors is also inconsistent with the SBIV theoretical structure.

Similar to the four-three-two-one model the four-two-one model at the highest level, is identified as an Abstract reasoning component. It is defined by same three components. This solution is rejected as all three components are from separate Reasoning Areas and add little to the interpretation of the SBIV.

The final two solutions examined at this age are the three-one and three-two-one models. Similar to the four-three-one model bipolar Verbal-Abstract/visual components are identified by both solutions. These solutions are rejected as the fail to meet the SBIV theoretical criteria and add little to understanding this test at the highest level. These results are displayed in Tables A.12 and A.13.

Ancillary Oblique Hierarchical Solutions for the 7 to 11 Years Subsample.

Tables A.14 to A.16 outline the four-three oblique hierarchical solution for the 7 to 11 years subsample. The first and third factors in this solution are characterized by *'g'*. The middle factor is identified as an Abstract/visual component. Congruence coefficients are weak. Therefore, this model is not able to add much to the interpretation of the SBIV.

The four-three-one model for this age group is displayed in Table A.17. As noted above this model is largely rejected at the second level. At the highest level an Abstract Reasoning factor is identified. This component is defined by Absurdities, Pattern Analysis and Bead Memory. All three variables are from different Reasoning Areas and such a grouping is inconsistent with the SBIV theoretical and test construction design.

Tables A.18 to A.20 display the four-three-two-one, four-two-one and three-two-one oblique solutions for the middle group. All three models arrive at a relatively similar solution. A strong General component is identified. Complexities at lower levels are largely responsible for their rejections.

TABLE A.0

SBIV (Four to) Three Oblique Factor (Pattern) 3 to 6 Years subsample

Variables	I	II	III
Vocabulary	7.791	-0.544	7.625
Comprehension	8.900	-1.786	7.422
Pattern Analysis	2.772	3.946	6.284
Copying	-3.250	8.959	5.829
Quantitative	5.140	1.942	6.517
Bead Memory	4.728	2.788	7.216
Memory for Sentences	5.626	1.938	7.839
Congruence Coefficients	.816	.852	.575

TABLE A.1

SBIV Four (to Three) Factor (Structure) Oblique Hierarchical Model

for the 3 to 6 Years Subsample

Variables	I	II	III
Vocabulary	-0.335	-0.363	0.415
Comprehension	-0.258	-0.300	0.340
Pattern Analysis	0.445	0.467	-0.390
Copying	-0. 138	-0.057	0.170
Quantitative	0.588	0.586	-0.523
Bead Memory	0.322	0.328	-0.254
Memory for Sentences	-0.245	-0.247	0.320

TABLE A.2

SBIV Four (To Three) Factor Oblique Hierarchical Correlation Coefficients

Factors	I	II	III
I	1.000		
II	.994	1.000	
III	-.995	-.992	1.000

TABLE A.3

SBIV Four-Three-Two Oblique Factor (Pattern) Hierarchical Model

for the 3 to 6 Years Subsample

Variables	I	II
Vcoabulary	.588	.181
Comprehension	.458	.128
Pattern Analysis	.247	.819
Copying	.887	.824
Quantitative	.054	.772
Bead Memory	.313	.725
Memory for Sentences	.676	.404
Congruence Coefficients	.634	.835

TABLE A.4

SBIV Four-Three-Two Oblique Factor (Structure) Hierarchical Model

for the 3 to 6 Years Subsample

Variables	I	II
Vocabulary	.515	-.057
Comprehension	.407	-.057
Pattern Analysis	-.084	.719
Copying	.554	.466
Quantitative	-.258	.750
Bead Memory	.020	.599
Memory for Sentences	.513	.131

TABLE A.5

SBIV Four-Three-Two Oblique Factor Intercorrelations

for the 3 to 6 Years Subsample

Factors	I	II
I	1.000	
II	-.404	1.000

TABLE A.6

SBIV Four-Three-One Oblique Factor Hierarchical Model

for the 3 to 6 Years Subsample

Variables	I
Vocabulary	-.372
Comprehension	-.300
Pattern Analysis	.435
Copying	.122
Quantitative	.567
Bead Memory	.302
Memory for Sentences	-.271

TABLE A.7

SBIV Four-Three Two One Oblique Factor Hierarchical Model

for the 3 to 6 Years Subsample

Variables	I
Vocabulary	.341
Comprehension	.277
Pattern Analysis	-.479
Copying	.053
Quantitative	-.602
Bead Memory	-.345
Memory for Sentences	.228

TABLE A.8

SBIV (Four-) Two Oblique Factor (Pattern) Hierarchical

for the 3 to 6 Years Subsample

Variables	I	II
Vocabulary	3.070	2.857
Comprehension	2.903	2.811
Pattern Analysis	2.678	2.979
Copying	2.890	2.435
Quantitative	2.629	3.148
Bead Memory	3.001	3.247
Memory for Sentences	3.290	3.056
Congruence Coefficients	.767	.642

TABLE A.9

SBIV (Four-) Two Oblique Factor (Structure) Hierarchical Solution

for the 3 to 6 Years Subsample

Variables	I	II
Vocabulary	.292	-.129
Comprehension	.169	-.012
Pattern Analysis	-.219	.374
Copying	.522	-.376
Quantitative	-.433	.592
Bead Memory	-.156	.328
Memory for Sentences	.318	-.144

TABLE A.10

SBIV (Four-) Two Oblique Factor Intercorrelations

for the 3 to 6 Years Subsample

Factors	I	II
I	1.000	
II	-.973	1.000

TABLE A.11

SBIV (Four-Two) One Oblique Factor Hierarchical Solution

for the 3 to 6 Years Subsample

Variables	I
Vocabulary	212
Comprehension	.091
Pattern Analysis	-.299
Copying	.452
Quantitative	-.516
Bead Memory	-.244
Memory for Sentences	.233

TABLE A.12

SBIV Three to One Factor Hierarchical Model for the 3 to 6 Years Subsample

Variables	I
Vocabulary	.425
Comprehension	.348
Pattern Analysis	-.444
Copying	-.074
Quantitative	-.553
Bead Memory	-.409
Memory for Sentences	.201

TABLE A.13

SBIV Three-Two-One Oblique Factor Hierarchical Model

for the 3 to 6 Years Subsample

Variables	I
Vocabulary	.450
Comprehension	.365
Pattern Analysis	-.412
Copying	-.030
Quantitative	-.531
Bead Memory	-.392
Memory for Sentences	.225

TABLE A.14

SBIV (Four to) Three Oblique Factor (Pattern) 7 to 11 Years Subsample

Variables	I	II	III
Vocabulary	1.044	.000	.915
Comprehension	.992	.033	.899
Absurdities	.632	.374	.699
Pattern Analysis	.538	.393	.733
Copying	.432	.924	-.251
Quantitative	.904	.185	.618
Bead Memory	.389	.163	.956
Memory for Sentences	1.049	-.205	1.021
Memory for Digits	1.032	-.119	.801
Memory for Objects	.687	.459	.498
Congruence Coefficients	.698	.766	.671

TABLE A.15

SBIV (Four to) Three Oblique Factor (Structure) 7 to 11 Years Subsample

Variables	I	II	III
Vocabulary	.496	.287	.289
Comprehension	.447	.321	.322
Absurdities	.142	.625	.519
Pattern Analysis	.023	.680	.620
Copying	.405	.708	-.018
Quantitative	.498	.340	.175
Bead Memory	-.216	.597	.810
Memory for Sentences	.476	.137	.284
Memory for Digits	.576	.108	.118
Memory for Objects	.300	.592	.330

TABLE A.16

SBIV (Four to) Three Oblique Factor Oblique

Hierarchical Correlation Coefficients

Factors	I	II	III
I	1.000		
II	-.192	1.000	
III	-.599	.533	1.000

TABLE A.17

SBIV Four-Three-One Oblique Factor Hierarchical Model

for the 7 to 11 Years Subsample

Variables	I
Vocabulary	-.048
Comprehension	-.095
Absurdities	-.423
Pattern Analysis	-.539
Copying	-.091
Quantitative	-.011
Bead Memory	-.695
Memory for Sentences	.003
Memory for Digits	.132
Memory for Objects	-.257

The three-one solution for the 7 to 11 years subsample is displayed in Table A.21. Similar to the four-three-one model on Abstract/visual component is identified at the highest level. Also like this latter model variables from both the Verbal and Memory Reasoning Areas are observed to load on this component. Therefore, it is inconsistent with basic SBIV tenets.

Table A.22 displays the two-one oblique solution for the middle subsample. A strong General component is found. This factor is supportive of the two component solution that identified both Crystallized and Fluid abilities. It is also consistent with one aspect of the SBIV theoretical structure.

<u>Oblique Hierarchical Solutions for the 12 to 23 Years Subsample</u>.

Tables A.23 to A.25 outline the four-three oblique hierarchical results. The first factor is largely *'g'*. The second factor is identified as an Abstract/visual component. Importantly, both Absurdities and Bead Memory are salients on this component. the third component is defined by salient loadings from both verbal and memory variables. Correlations between the second and third components are moderate. Despite these relatively clear results the congruence coefficients are small. This is largely due to the failure of the verbal and memory variables to form separate factors, at this level.

Tables A.26 to A.28 display the four-three-two oblique hierarchical results. The first factor is identified as *'g'*. It is defined by salient loadings from all variables with the exception of Copying. The congruence coefficient is only moderate for this factor.

The second component is largely Abstract/visual. Importantly, Absurdities, Bead Memory and Memory for Objects are observed to load on this factor. Similar to the first factor the congruence coefficient is moderate.

Tables A.29 to A.31 displays the four-two oblique solutions for the oldest subsample. Both General and Abstract/visual factors are identified. These results are observed to closely parallel those obtained in the four-three-two solution. They are viewed as largely unsatisfactory because of their lack of interpretative value.

TABLE A.18

SBIV Four-Three Two One Oblique Factor Hierarchical Model

for the 7 to 11 Years Subsample

Variables	I
Vocabulary	-.571
Comprehension	-.581
Pattern Analysis	-.716
Copying	-.864
Quantitative	-.606
Bead Memory	-.536
Memory for Sentences	-.424
Memory for Digits	-.430
Memory for Objects	-.748

TABLE A.19

SBIV (Four-Two) One Oblique Factor Hierarchical Solution

for the 7 to 11 Years Subsample

Variables	I
Vocabulary	-.814
Comprehension	-.810
Absurdities	-.817
Pattern Analysis	-.817
Copying	-.622
Quantitative	-.736
Bead Memory	-.733
Memory for Sentences	-.735
Memory for Digits	-.671
Memory for Objects	-.786

TABLE A.20

SBIV Three-Two-One Oblique Factor Pattern Hierarchical Model

for the 7 to 11 Years Subsample

Variables	I
Vocabulary	.768
Comprehension	.775
Absurdities	.762
Pattern Analysis	.840
Copying	.678
Quantitative	.756
Bead Memory	.758
Memory for Sentences	.744
Memory for Digits	.709
Memory for Objects	.789

TABLE A.21

SBIV Three to One Hierarchical Model

for the 7 to 11 Years Subsample

Variables	I
Vocabulary	.274
Comprehension	.293
Absurdities	.669
Pattern Analysis	.578
Copying	.293
Quantitative	.138
Bead Memory	.608
Memory for Sentences	.043
Memory for Digits	-.104
Memory for Objects	.433

TABEL A.22

SBIV Two to One Oblique Factor Pattern Hierarchical Model

for the 7 to 11 Years Subsample

Variables	I
Vocabulary	.779
Comprehension	.784
Absurdities	.748
Pattern Analysis	.830
Copying	.674
Quantitative	.769
Bead Memory	.745
Memory for Sentences	.765
Memory for Digits	.735
Memory for Objects	.786

TABLE A.23

SBIV (Four to) Three Oblique Factor (Pattern) 12 to 23 Years Subsample

Variables	I	II	III
Vocabulary	.615	.043	.557
Comprehension	.732	.011	.457
Absurdities	.452	.317	.413
Pattern Analysis	.227	.729	.156
Matrices	.560	.393	.222
Quantitative	.840	.077	.199
Number Series	.665	.285	.295
Bead Memory	.382	.564	.139
Memory for Sentences	.813	-.240	.582
Memory for Digits	.693	.051	.318
Memory for Objects	.185	.006	.887
Congruence Coefficients	.743	.711	.659

TABLE A.24

SBIV (Four to) Three Oblique Factor (Structure) 12 to 23 Years Subsample

Variables	I	II	III
Vocabulary	.692	.580	.639
Comprehension	.785	.560	.537
Absurdities	.643	.714	.602
Pattern Analysis	.585	.906	.506
Matrices	.770	.761	.458
Quantitative	.897	.561	.319
Number Series	.830	.730	.491
Bead Memory	.661	.805	.430
Memory for Sentences	.760	.402	.559
Memory for Digits	.750	.519	.413
Memory for Objects	.279	.489	.908

TABLE A.25

SBIV (Four to Three) Factor Oblique Hierarchical Correlation Coefficients

for the 12 to 23 Years Subsample

Factors	I	II	III
I	1.000		
II	.470	1.000	
III	.130	.477	1.000

TABLE A.26

SBIV Four-Three-Two Oblique Factor (Pattern) Hierarchical Model

for the 12 to 23 Years Subsample

Variables	I	II
Vocabulary	.543	.442
Comprehension	.666	.297
Absurdities	.561	.467
Pattern Analysis	.615	.457
Matrices	.756	.285
Quantitative	.865	.039
Number Series	.785	.279
Bead Memory	.683	.322
Memory for Sentences	.581	.285
Memory for Digits	.676	.184
Memory for Objects	.016	.865
Congruence Coefficients	.822	.711

TABLE A.27

SBIV Four-Three-Two Oblique Factor (Structure) Hierarchical Model

for the 12 to 23 Years Subsample

Variables	I	II
Vocabulary	.729	.670
Comprehension	.791	.577
Absurdities	.757	.703
Pattern Analysis	.807	.715
Matrices	.875	.602
Quantitative	.882	.402
Number Series	.902	.609
Bead Memory	.819	.609
Memory for Sentences	.700	.529
Memory for Digits	.753	.468
Memory for Objects	.379	.872

TABLE A.28

SBIV Four-Three-Two Oblique Factor Intercorrelations

Factors	I	II
I	1.000	
II	.420	1.000

TABLE A.29

SBIV (Four-) Two Oblique Factor Pattern Hierarchical Solution

for the 12 to 23 Years Subsample

Variables	I	II
Vocabulary	.564	.389
Comprehension	.744	.201
Absurdities	.471	.482
Pattern Analysis	.371	.565
Matrices	.697	.269
Quantitative	.991	-.098
Number Series	.778	.230
Bead Memory	.539	.365
Memory for Sentences	.745	.140
Memory for Digits	.756	.087
Memory for Objects	-.108	.933
Congruence Coefficients	.877	.772

TABLE A.30

SBIV (Four-) Two Oblique Factor Structure Hierarchical Solution

for the 12 to 23 Years Subsample

Variables	I	II
Vocabulary	.831	.776
Comprehension	.882	.711
Absurdities	.801	.805
Pattern Analysis	.759	.820
Matrices	.881	.747
Quantitative	.924	.582
Number Series	.935	.763
Bead Memory	.789	.734
Memory for Sentences	.841	.651
Memory for Digits	.816	.605
Memory for Objects	.531	.859

TABLE A.31

SBIV (Four-) Two Oblique Factor Intercorrelations

for the 12 to 23 Years Subsample

Factors	I	II
I	1.000	
II	.685	1.000

Tables A.32 to A.34 display the four-three-one, four-three-two-one and four-two-one oblique factor solutions for the oldest subsample. All solutions result in a strong *'g'* component. These results are, therefore, supportive of *'g'*. However, these models are rejected because of their less than satisfactory results of their intermediate levels.

Tables A.35 to A.38 provide the three-two oblique hierarchical results for the 12 to 23 years subsample. Similar to the four-three-two and four-two results the first component is largely *'g'* and the second is Abstract/visual. Absurdities and Bead Memory are observed to load in a salient fashion on this second factor. This is also consistent with the previously identified solutions. The three-two oblique hierarchical model is rejected at this level given its failure to conform to any of the SBIV substructures.

Table A.39 provides the three-one oblique hierarchical results for the 12 to 23 subsample. A strong *'g'* component is found. These results are supportive of the highest level of the developmental that is proposed for this age-group in Chapter 5.

The two-one oblique hierarchical solution for the oldest group is displayed in Table A.40. This solution is complex and largely unidentifiable. Salient loadings are observed from Verbal, Abstract/visual and Quantitative Reasoning Areas. Consequently, this solution has little explanatory value.

Summation

A variety of hierarchical solutions are presented and subsequently rejected. These results suggest that very few hierarchical arrangements adhere to the SBIV structure or substructures. The equivocal nature of these results questions the robustness of the SBIV hierarchical model and its various substructures.

TABLE A.32

SBIV Four-Three-One Oblique Factor Hierarchical Model

for the 12 to 23 Year Subsample

Variables	I
Vocabulary	.832
Comprehension	.819
Absurdities	.868
Pattern Analysis	.906
Matrices	.887
Quantitative	.780
Number Series	.908
Bead Memory	.855
Memory for Sentences	.736
Memory for Digits	.735
Memory for Objects	.722

TABLE A.33

SBIV Four-Three Two One Oblique Factor Hierarchical Model

for the 12 to 23 Year Subsample

Variables	I
Vocabulary	-.830
Comprehension	-.812
Absurdities	-.866
Pattern Analysis	-.903
Matrices	-.877
Quantitative	-.762
Number Series	-.897
Bead Memory	-.848
Memory for Sentences	-.729
Memory for Digits	-.724
Memory for Objects	-.743

TABLE A.34

SBIV Four-Two-One Oblique Factor Hierarchical Solution

for the 12 to 23 Years Subsample

Variables	I
Vocabulary	.875
Comprehension	.868
Absurdities	.875
Pattern Analysis	.860
Matrices	.887
Quantitative	.820
Number Series	.925
Bead Memory	.830
Memory for Sentences	.813
Memory for Digits	.774
Memory for Objects	.757

TABLE A.36

SBIV Three to One Oblique Factor (Structure) Hierarchical Model

for the 12 to 23 Years Subsample

Variables	I	II
Vocabulary	.861	.659
Comprehension	.863	.663
Absurdities	.758	.767
Pattern Analysis	.629	.922
Matrices	.781	.836
Quantitative	.822	.690
Number Series	.852	.822
Bead Memory	.689	.851
Memory for Sentences	.910	.518
Memory for Digits	.823	.619
Memory for Objects	.787	.483

TABLE A.37

SBIV Three to Two Correlations Between Factors

for the 12 to 23 Years Subsample

Factors	I	II
I	1.000	
II	.597	1.000

TABLE A.38

SBIV Three-Two-One Oblique Factor Hierarchical Model

for the 12 to 23 Years Subsample

Variables	I
Vocabulary	-.850
Comprehension	-.854
Absurdities	-.853
Pattern Analysis	-.868
Matrices	-.905
Quantitative	-.846
Number Series	-.937
Bead Memory	-.862
Memory for Sentences	-.799
Memory for Digits	-.807
Memory for Objects	-.711

TABLE A.39

SBIV Three to One Factor Hierarchical Model

for the 12 to 23 Years Subsample

Variables	I
Vocabulary	.855
Comprehension	.858
Absurdities	.853
Pattern Analysis	.861
Matrices	.903
Quantitative	.849
Number Series	.937
Bead Memory	.858
Memory for Sentences	.809
Memory for Digits	.812
Memory for Objects	.718

TABLE A.40

SBIV Two to One Oblique Factor Hierarchical Model

for the 12 to 23 Years Subsample

Variables	I
Vocabulary	.248
Comprehension	.348
Absurdities	.434
Pattern Analysis	.459
Matrices	.354
Quantitative	.348
Number Series	.364
Bead Memory	.227
Memory for Sentences	.081
Memory for Digits	-.073
Memory for Objects	-.385