THE APPROPRIATENESS OF SELECTED SUBTESTS OF THE STANFORD-BINET INTELLIGENCE SCALE: FOURTH EDITION FOR HEARING IMPAIRED CHILDREN

by

JO-ANNE PERLEY-MCFIELD

B.P.E., The University of British Columbia, 1973

Dip. Education of the Hearing Impaired, The University of British Columbia, 1984

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

(Department of Educational Psychology and Special Education)

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

June 1990

(Department of Educational Psychology and Special Education)

The University of British Columbia
2075 Wesbrook Place
Vancouver, Canada
V6T 1W5

Date: June 1990

ABSTRACT

This study proposed to evaluate the appropriateness of selected subtests of the Stanford-Binet Intelligence Scale: Fourth Edition (SB:FE) for use with severely to profoundly hearing impaired children. The subjects used in this study were enrolled in a residential/day school for the deaf whose educational methodology was Total Communication. The subjects were tested on both the SB:FE nonverbal selected subtests and the Performance Scale of the Wechsler Intelligence Scale for Children-Revised (WISC-R PIQ).

To assess appropriateness, several procedures were employed comparing data gathered from the hearing impaired sample with data reported for the standardized population of the SB:FE. Correlations were computed between the WISC-R and the SB:FE and comparisons of the total composite scores for each measure were made to detect any systematic differences.

The results indicated that the correlations reported for the hearing impaired sample are generally similar to the correlations reported for the standardized sample of the SB:FE. The analysis performed between the Area Scores of the SB:FE and the WISC-R PIQ to detect systematic differences revealed a difference of one standard deviation between these two instruments, with the SB:FE results being lower than the WISC-R PIQ results.

It was concluded that the selected subtests of the SB:FE and the WISC-R PIQ could not be used interchangeably. Further research into this area was advised before using this measure to estimate general cognitive ability for hearing impaired children whose levels of language development may be delayed. Further research was also encouraged to confirm the suggestion of greater predictive validity of the SB:FE with academic measures. It was suggested that these

findings indicated that the use of language as a cognitive tool may be important in acquiring certain problem solving skills.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I.

A. BACKGROUND TO THE PROBLEM

It is possible to trace the history of intelligence testing of the hearing impaired and see it as a journey of discovery of the many confounding variables involved in any research. Just as the concept of intelligence and its measurement cannot be simply explained, neither is there a simple condition referred to as hearing impairment. Research based on different interpretations of these many variables has produced conflicting results and perpetuated many misconceptions.

Pintner, an early investigator in this field, produced findings from which he concluded the deaf were intellectually inferior to the hearing (Pintner, Eisenson & Stanton, 1941). The fact that language measures were used to form conclusions about intelligence for this population did not, at first, appear to be inappropriate. Over the next few decades however, the use of non-verbal, performance type instruments, individually administered, came to be regarded as more appropriate for the assessment of hearing impaired children.

A series of studies by Myklebust (1960) demonstrated that while the hearing impaired achieved global scores within the average range for hearing children, there appeared to be characteristic subtest variations. These variations were interpreted as indicating the hearing impaired population were limited in their ability for abstract reasoning. Myklebust hypothesized this was due to the sensory impairment which caused a different cognitive structuring in the brain.

Yet another line of research, led by Rosenstein (1961), Furth (1966) and Vernon (1967) began to indicate that hearing impaired and hearing people were very similar in their cognitive functioning. The number of studies in the last two decades investigating the nature of cognitive functioning of the hearing impaired indicates that this area continues to intrigue researchers.

The search for the most appropriate instrument to assess intelligence and appropriate administration procedures has concerned educators and psychologists working with the hearing impaired. The basic principles to emerge have been (a) that the instrument must be individually administered and be a non-verbal performance test; (b) that the administration procedures must ensure that the child understands the exact nature of the task; (c) that the instrument must have respected psychometric properties concerning validity and reliability as well as a wide research base to validate its use as a diagnostic tool and (d) that plurality of instrumentation is necessary for appropriate assessment (Sullivan and Vernon, 1979).

The Performance Scale of the Wechsler Intelligence Scale for Children — Revised (WISC-R PIQ) (Wechsler, 1974) has emerged as the leading assessment instrument for testing hearing impaired children in North America (Allen, Abraham, & Stoker, 1988; Levine, 1974) and in Great Britain (Kyle, 1980). Professionals should not, however, become complacent with this situation. As new intelligence measures appear on the market, their use with the hearing impaired population should be explored. The need for plurality of instrumentation has been considered vital to thorough and appropriate assessment practice (Sullivan and Vernon, 1979; Sattler, 1982), as well as contributing to our knowledge of cognitive functioning.

The Stanford-Binet Intelligence Scale: Fourth Edition (SB:FE) (Thorndike, Hagen & Sattler, 1986b) has been published. Its suitability for use with the hearing impaired population needs to be explored. The previous edition of the

Stanford-Binet Intelligence Scale (Terman and Merrill, 1973) was considered totally inappropriate for the assessment of intelligence in the hearing impaired population because of its heavy emphasis on language ability (Sullivan and Vernon, 1979). The fourth edition represents a major revision of the scale. The authors decided that four areas of cognitive ability should be appraised: verbal reasoning, quantitative reasoning, abstract/visual reasoning, and short-term memory. It was also decided that the revised scale would continue to provide an overall score that would represent general reasoning ability.

On the surface, the new changes to the Stanford-Binet Intelligence Scale appear to make its use with the hearing impaired population potentially appropriate. The separation of the Verbal Reasoning Area from the Quantitative Reasoning Area, the Abstract/Visual Reasoning Area and the Short-Term Memory Area suggests that these latter areas may not rely heavily on verbal abilities. Many subtests of these areas appear to be essentially nonverbal, performance-type tests. As such their use with hearing impaired children may be appropriate.

Within the Short-Term Memory Area of this fourth edition of the Stanford-Binet, the division of visual memory into two separate subtests may yield particularly useful information. Studies by Watson, Sullivan, Moeller and Jensen (1982) have reported high correlations between visual memory and English language ability.

For the older child especially, the Stanford-Binet Intelligence Scale: Fourth Edition provides a wide range of subtests, sampling various areas of nonverbal abilities. A total of eight subtests may be used with the older child, providing more diagnostic information and increasing the reliability of the intellectual estimate.

Yet another attractive feature of this instrument is that it functions as a power test rather than a test rewarding fast responses. The lack of timed responses is regarded as especially important in the assessment of hearing impaired children (Sullivan and Vernon, 1979). The one timed subtest, Pattern Analysis, only sets a time limit and does not reward for extra speed.

In addition to these promising features, the new edition of the Stanford-Binet has other attractions for use with a hearing impaired population. Each measuring instrument provides to some degree a unique perspective of the construct that it purports to measure. The perspective offered by the Stanford-Binet Intelligence Scale: Fourth Edition may well provide new information to the field of cognitive functioning and reveal exciting new areas for investigation.

B. PURPOSE OF THE STUDY

The purpose of this study was to evaluate selected subtests of the new Stanford-Binet Intelligence Scale: Fourth Edition (SB:FE) on its appropriateness for use with hearing impaired children whose educational methodology is Total Communication (TC). To assess appropriateness, several procedures were employed. The patterns of intercorrelations between the Performance Scale of the Wechsler Intelligence Scale for Children — Revised and the selected subtests of the SB:FE for the sample of hearing impaired children were compared with similar correlations reported for the non-exceptional sample of the SB:FE. Further evidence of the construct validity of the new SB:FE for the hearing impaired population was sought by comparing total scores from these two intelligence measures. Evidence of appropriate internal consistency was sought by conducting exploratory item analysis on the SB:FE results. The predictive validity was also explored by correlating this instrument with academic achievement measures available from results of the Stanford Achievement Tests – Hearing Impaired (SAT-HI) (Gardner, Rudman, Karlsen & Merwin, 1982).

C. HYPOTHESES

The hypotheses to be tested in this study:

1. For the hearing impaired population in this study, the pattern of correlations between the Performance Scale of the WISC-R and the Abstract/Visual, Quantitative, and Short-Term Memory Area scores of the Stanford-Binet: Fourth Edition will not be significantly different from that reported for the non-exceptional sample.
2. For the hearing impaired population in this study, the pattern of subtest intercorrelations among the SB:FE will not be significantly different from that reported for the non-exceptional sample.
3. For the hearing impaired population in this study, the pattern of subtest intercorrelations among the WISC-R PIQ will not be significantly different from that reported for the standardization sample.

Exploratory research questions to be addressed:

1. Exploratory item analysis will be conducted to examine item characteristics of the various subtests of the SB:FE to determine significant trends for this sample of hearing impaired children.
2. For the hearing impaired population in this study, will either instrument systematically yield greater composite scores than the other?
3. What is the pattern of subtest intercorrelations between these two

instruments? For instance, are certain subtests characteristically associated on both instruments, e.g., Block Design and Pattern Analysis?

4. What are the correlations produced between the WISC-R PIQ, the SB:FE and the SAT-HI academic measures?

D. DEFINITION OF TERMS

Throughout the study the following terms will be used as defined below:

Deaf

A term indicating a hearing disability which precludes successful processing of linguistic information through audition, with or without a hearing aid. A deaf individual would have a hearing loss of 90dB or greater in the better unaided ear across the speech frequencies (500, 1000, and 2000 Hz).

Hearing Impairment

A term used to describe a person who has any degree of hearing loss. It is a generic term which includes both deaf and hard of hearing persons.

American Sign Language (ASL)

American Sign Language (ASL) is the language of the signing adult deaf community in North America (Baker and Cokely, 1980).

Pidgin Sign English

A sign communication that incorporates grammar from both ASL and English. The syntax can vary from being more like ASL to being more like English (Woodward, 1973).

Nonverbal Performance Tests

These are tests that do not require verbal responses. Tasks may be performed by manipulation of materials or by indicating choices using a pointing response. These tests may require verbal directions.

Total Communication

A philosophy requiring the incorporation of appropriate aural, manual, and oral modes of communication with and among hearing impaired persons. Definition accepted by the Conference of Executives of American Schools for the Deaf in 1976 (Gannon, 1981, p. 369).

CHAPTER II. SURVEY OF THE LITERATURE

This chapter has been organized to present three major topics. The first provides a brief historical overview of intelligence testing with the hearing impaired. The second topic addresses the issue of which testing instruments are in current use and reviews the findings of previous research with the WISC-R PIQ. The third topic reviews the modifications used in test instructions for the hearing impaired and presents the rationale for the administration modifications used in this study.

A. HISTORICAL OVERVIEW

There has been considerable research interest in intelligence testing with the hearing impaired during this century. In general, the thrust of research inquiry has been toward establishing the relative intelligence of deaf and hearing children. Researchers have evolved from the position of regarding the deaf as intellectually inferior, to the present day position which considers that while some differences in test performance may be observed, in general the deaf as a group score within the average range of intelligence.

What is apparent in reviewing the research in this area is that much of the inconsistencies in the findings have been a result of the different variables intervening into the testing situation and the different measuring instruments used. As experimental control has increased, the observed differences between the deaf and hearing have decreased.

The early research conclusions reached by Pintner and his colleagues illustrate the significance of type of measuring instrument used. Pintner, Eisenson and Stanton (1941) concluded that deaf people were intellectually inferior to hearing people and showed definite deficits in various aspects of cognitive functioning. What is important to note is that most of Pintner's tests required the use of verbal English language both in administration and in response. The use of a verbal measure to assess the reasoning abilities of a deaf child are considered inappropriate today (Sattler, 1982; Sullivan and Vernon, 1979).

Vernon's (1968) summary of the research on the intelligence of the hearing impaired also contributed importantly to the identification of intervening variables in the research process. After reviewing fifty years of research in this area, Vernon concluded that when there are no complicating multiple handicaps, the hearing impaired function at approximately the same IQ level on performance intelligence tests as do the hearing. Vernon had examined closely the relationship of etiology of deafness to intelligence. A disproportionately higher prevalence of low IQs for the hearing impaired population had been noticed. Vernon suggested that many of the etiologies of profound hearing loss were also responsible for other neurological impairment which would result in lower intelligence. Sullivan (1982) has suggested that the distribution of intelligence in hearing impaired children may be bimodal. The multihandicapped group would tend to form a lower group and non-multihandicapped would form a higher group.

However, despite advances in experimental control and improved assessment procedures, not all the observed differences between deaf and hearing children disappeared. There still were enough inconsistencies in the findings to intrigue researchers. There is disagreement as to how these observed differences should be interpreted. If these observed differences are in fact real, it would be logical to conclude that the deaf have cognitive differences from the hearing. Myklebust (1966) represented this view and proposed a theory of sensory deprivation to explain these differences. Myklebust found that while the hearing impaired performed within the average range of general cognitive ability, there were certain subtest variations that indicated a qualitative difference in mental functioning. He concluded that the deaf fell below average mainly on tests which required a type of abstraction and reasoning process. To account for these observed variations in the mental abilities of deaf children, Myklebust proposed the organismic shift hypothesis. This stated that when a sensory deprivation is present from early life, the organism must modify its means for maintaining adequate environmental contact. A shift in perceptual organization was made to maintain psychological equilibrium. This resulted in a different cognitive structuring from normally hearing children.

Levine (1981) suggested that the cognitive differences observed in the deaf were due to different experiences. This experiential defecit theory suggested that the environment has a powerful influence in the fashioning of human behavior. The absence of sound from the environment of the hearing impaired would create psychological voids.

Researchers such as Rosenstein (1961), Furth (1966) and Vernon (1967) have supported the position that deaf and hearing impaired people are cognitively similar in all important abilities. Observed differences in test results may be due to a remaining lack of experimental control over verbal and other factors that influence performance on tasks that are believed to be nonverbal. Research has been done that shows the method of administration of tests has significant effects on test results. This issue will be discussed in more detail in a subsequent section.

The validity of using Performance type tests as an estimate of intelligence has been the subject of further research. When Pintner produced his research on intelligence (using language measures) the current theoretical position favoured the belief that language was primary and thinking took place in language. Therefore the use of language to assess intelligence was logical and acceptable at that time. As the research in the field of intelligence accumulated, different interpretations of intelligence began to evolve. A position known as the cognitive dominant hypothesis began to receive research support. This position proposed that basic perceptual and cognitive development can proceed independently from language. The acceptance that all cognitive processes do not necessarily need language tended to support the use of performance scales to assess cognitive ability.

Very few research studies investigating the predictive validity of intelligence measures with the hearing impaired school population have been reported. Research into the relationship of nonverbal intelligence measures and academic achievement in the hearing population suggests this relationship is not particularly strong. Zimmerman and Woo-Sam (1972), in a series of studies relating the WISC (Wechsler, 1949) to various achievement measures, reported a median correlation of .33 between Performance IQ and achievement. Murphy (1957) began to investigate this issue in the hearing impaired population and found WISC Performance Scale results to correlate 0.64 with teacher's ratings of current academic attainment. This rather strong coefficient has not been supported, however, in studies using standardized test instruments. Brill (1962) conducted a study using the Wechsler Adult Intelligence Scale (WAIS) and the WISC (Wechsler, 1949). Correlations of .54 and .55 were found between the nonverbal intelligence measures and two standardized achievement measures.

Hirshoren, Hurley and Kavale (1979) correlated the WISC-R PIQ with the Stanford Achievement Test. The most robust single correlation of .35 was found between the Average Grade Score and the WISC-R PIQ. Brooks and Riggs (1980) correlated the WISC and WISC-R Performance IQs with reading achievement using the SAT-HI as well as class observations and the results were found to be nonsignificant. In 1982, Watson, Sullivan, Moeller and Jensen explored the relationship between nonverbal intelligence and English language ability in deaf children. Using both the WISC-R PIQ and the Hiskey-Nebraska Test of Learning Abilities (H-NTLA) (Hiskey, 1966), they found average correlations of .45 between these measures and language measures. They found further, that subtests requiring visual memory contributed most significantly to the multiple regression analysis which yielded an average correlation of .68. A further study by Watson, Goldgar, Kroese and Lotz (1986) found a median correlation of .42 between the H-NTLA and various measures of academic achievement. Correlations between the WISC-R PIQ and these academic measures were lower with a median of .24. The authors suggested that the strength of the H-NTLA achievement correlations may be due to the heavier weighting this test gives to visual memory tasks. They suggested that visual memory skills may be particularly important in the hearing impaired child's acquisition of language and knowledge.

These rather low reported correlations between performance measures and academic measures do not appear to provide much support for the use of these intelligence measures as an indicator of academic aptitude. However, it should be considered that these measures do provide some information. If used in conjunction with other assessment instruments, useful information can be provided to assist in educational planning. It is most important however, for those involved with such decision making, to be well aware of the limitations of these test instruments.

B. USE OF THE PERFORMANCE SCALE OF THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN-REVISED

The Performance Scale of the WISC-R has been used as the comparison measure in this study. This was done to parallel studies conducted on the construct validity of the SB:FE. The data for these studies are available in the Technical Manual of the SB:FE (Thorndike, Hagen & Sattler, 1986a). The studies were carried out to determine the correlations between the SB:FE and several individual intelligence tests. These studies were conducted using both non-exceptional samples and special groups of children. Of the individual intelligence tests used in these construct validity studies, the WISC-R PIQ was deemed the most appropriate for use in this study as the comparison measure. Evidence will be presented to show that it is a nonverbal performance type of instrument, considered most suitable for testing the hearing impaired. It has been found to be the most frequently used intelligence measure with the hearing impaired. Further, an extensive body of research has accumulated on the use of this instrument with hearing impaired populations.

It has been noted that most observed differences in intelligence tests results between deaf and hearing have decreased as experimental control has increased. The identification of the many confounding variables involved in research work with the deaf has also led to improved assessment procedures for educators, pyschologists, and other professionals working with the deaf.

One of the first principles to be formulated defined the type of test used in assessing intelligence in deaf children. Lane (1938), a psychologist working with the deaf presented an early paper on the measurement of mental and educational ability of the deaf children which indicates the stage of psychological testing practices of that period. Lane acknowledged that the issue of whether a performance test could replace a linguistic test was still controversial, and that the problem of comparing the intelligence of hearing and hearing impaired depended on the definition of intelligence accepted by the investigator. Her recommendations for valid psychological measurement included the use of nonverbal performance type instruments. Lane (1938) observed that

> . . . any test requiring linguistic responses or the understanding of verbal instructions becomes a test of the educational achievement and lipreading ability of the deaf as well as a measure of native intelligence. Selection must therefore be made from those performance tests in which the directions are in pantomime and the responses are manipulative (p. 169).

She further recommended that individual tests were to be preferred over group tests because of the influence of such external factors as attention to the examiner, understanding of directions, and time factors.

The use of individually administered, nonverbal performance tests continues to be recommended to the present day. Commonly used intelligence measures standardized on hearing impaired school children which meet these requirements include:

1. the Hiskey-Nebraska Test of Learning Aptitude (H-NTLA) (Hiskey, 1966);
2. the Non-Verbal Intelligence Tests for Deaf and Hearing Subjects (Snidjers and Snidjers-Oomen, 1970);
3. the Ontario School Ability Examination (Amoss, 1949); and

4. the Wechsler Intelligence Scale for Children-Revised, Performance Scale, Deaf Norms (Anderson and Sisco, 1977).

Commonly used intelligence measures not standardized on hearing impaired school children which also meet the requirements are:

1. the Arthur Adaptation of the Leiter International Performance Scale (Arthur, 1955);
2. the Columbia Mental Maturity Scale (Burgemeister, Blum & Lorge, 1972);
3. the Goodenough-Harris Draw-a-Man (Goodenough & Harris, 1963);
4. the Leiter International Performance Scale (Leiter, 1948);
5. the Standard Progressive Matrices (Raven, 1960); and
6. the Wechsler Intelligence Scale for Children-Revised, Performance Scale, (Wechsler, 1974).

These lists are not intended to be exhaustive but mention those tests most commonly in use with the hearing impaired school aged child (Allen *et al.*, 1988; Bragman, 1982; Kyle, 1980; Levine, 1971; Levine, 1974; Sullivan and Vernon, 1979; Vernon, 1976; Vernon and Brown, 1964).

Of these tests, the Performance Scale of the Wechsler Intelligence Scale for Children – Revised (Wechsler 1974) has been reported as the most commonly used in psychological assessment of school age hearing impaired children. The deaf norms are not reported to be widely in use. These norms were produced from a study conducted by Anderson and Sisco (1977) on a national sample of 1,228 deaf children. These children were from 18 residential schools and 4 day-schools for deaf children located throughout the United States. A major concern regarding the use of these norms appears to be the lack of standardized administration procedures.

In a United States national survey of psychologists working with the hearing impaired, the WISC-R Performance Scale was reported as the most frequently used intelligence measure by an overwhelming majority of psychologists (Levine, 1974). Another survey by Spragins and Gibbins (1982), though not national in scope, indicates that the popularity of this instrument has continued. In their survey of school psychologists working with hearing impaired children, they found that the WISC-R was used by 91% of the respondents. In New York and the New England states, a survey indicated that the WISC-R Performance Scale was still ranked the most popular test in use for hearing impaired children (McQuaid and Alovisetti, 1981). That the WISC-R PIQ still maintains this position was further demonstrated by the results of the Allen *et al.* (1988) survey of psychoeducational assessment of the hearing impaired in both the United States and Canada. In the United Kingdom too, the Performance Scale of the WISC-R has been cited as being the most widely used for hearing impaired children (Kyle, 1980).

In view of the small number of intelligence tests standardized on hearing impaired children, non-deaf standardized tests have maintained an important position in psychological test practices. Leading authorities in the field of psycho-educational assessment have maintained various positions on the issue of whether a test developed and standardized on a hearing population can be considered suitable for use with a hearing impaired population.

Gerweck and Ysseldyke (1975) and later Salvia and Ysseldyke (1985) maintain a firm position against the use of tests standardized on a hearing population. They have stated that their use with a hearing impaired population violates the basic assumptions underlying psychological assessment. Anastasi

(1982) also criticizes modifications used in standardized testing procedures. She states that one cannot assume that reliability, validity and norms remain unchanged in these circumstances. Both Anastasi (1982) and Salvia and Ysseldyke (1985) recommend only the Hiskey-Nebraska Test of Learning Aptitude (Hiskey, 1966) for use with hearing impaired school children.

Sattler (1982) specifically recommends the Performance Scale of the WISC-R as a reliable and valid instrument for assessing the intelligence of deaf children and asserts that the 1974 published norms are valid for the assessment of deaf children. He further claims there is no evidence that intelligence, as measured by the WISC-R Performance Scale, differs qualitatively in deaf and hearing groups. A study by Hirshoren, Kavale, Hurley and Hunt (1977) reported that with respect to both individual subjects and total test reliability, the Performance Scale of the WISC-R possessed a satisfactory level of reliability for testing the mental abilities of a deaf population. Reliability coefficients were compared for each subtest and for the composite Performance Scale of the WISC-R. For the Object Assembly, Block Design and Picture Completion subjects, the reliability values did not differ significantly from the standardization sample. One subtest, Picture Arrangement, was found to have a higher reliability coefficient. It was suggested that the greater variability found in the deaf sample on this subtest may have contributed to this finding. The composite reliability of the WISC-R was based on the reported reliabilities of four subtests. The composite total ($r=.90$) is comparable to the total Performance Scale reliability reported by Wechsler (1974). Braden (1985) reports that the WISC-R Performance Scale has the same factor structure in deaf and hearing groups. Braden concluded that the internal validity, construct validity, and criterion validity of the

WISC-R Performance Scale are essentially the same for deaf children as for hearing children.

Vernon has written extensively in the field of psychological assessment of the hearing impaired. Vernon includes the Performance Scale of the WISC-R in his published lists of recommended tests and reviewed as an excellent test for use with hearing impaired children (Sullivan & Vernon, 1979; Vernon and Brown, 1964; Vernon, 1976). The issues of the lack of standardized administration procedures is not ignored, however. Although test administration modifications are advocated, their use is noted to affect test reliability by introducing sources of inconsistency and so underscores the importance of administering more than one performance intelligence measure.

The research information which has accumulated on the use of the WISC-R Performance Scale also contributes to its desirability for use as a comparison measure in this study. Much of this research provides information that supports its use with hearing impaired children. There have been a limited number of studies exploring the correlations between the WISC-R PIQ and other nonverbal intelligence measures. Table 1 displays results from some of these studies.

Other studies report mean score same group comparisons of the WISC-R PIQ with other nonverbal intelligence measures. The results of these studies, presented in Table 2, do not yield consistent findings as to whether the WISC-R PIQ generally scores higher or lower than other measures.

In summary, the accumulated evidence regarding the use of the WISC-R Performance Scale in the assessment of hearing impaired children is regarded as substantial enough to support its value as a comparison measure in this study.

Table 1

WISC-R PIQ Concurrent Validity Correlations

Study	Subjects	Criterion	r
Ritter (1976)	31 children $\bar{x}$ age=6.9 yr. HL mild to moderate	AA-LIPS CPM	.78 .50
Hirshoren, Hurley & Kavale (1979)	59 children $\bar{x}$ age=10.6 yr. prelingually deaf	H-NTLA	.89
Evans (1980)	125 5-12 yr. prelingually deaf		
	ages 5-6	CPM	.56
	ages 7-8	CPM	.56
	ages 9-10	CPM	.65
	ages 11-12	CPM	.83
Ullissi & Gibbons (1984)	40 children $\bar{x}$ age 8 yr. prelingually deaf	LIPS	.82
James (1984)	34 children age 6-11 HL $\bar{x}$=85dB	CPM	.87
	50 children age 11-16.5 yr. HL $\bar{x}$=80dB	SPM	.78
Watson & Goldgar (1985)	71 children $\bar{x}$ age=13.2 yr. HL profound 44 HL severe 22 HL mod. severe 5	H-NTLA	.85
Phelps & Ensor (1986)	49 children $\bar{x}$ age=11.51 yr. HL-prelingually deaf (moderate to profound)	H-NTLA	.91 WISC-R deaf norms
Porter & Kirby (1986)	49 children ages 7-12 yr. $\bar{x}$ age=10.3 yr. HL $\bar{x}$=102dB	KABC non-verbal scale	
	• pantomine gesture group		.68
	• ASL group		.64

Table 1 continued

AALIPS: Arthur Adaptation of the Leiter International Performance Scale (Arthur, 1955)
CPM: Raven's Colored Progressive Matrices (Raven, 1965)
H-NTLA: Hiskey-Nebraska Test of Learning Aptitude (Hiskey, 1966)
LIPS: Leiter International Performance Scale (Leiter, 1948)
SPM: Standard Progressive Matrices (Raven, 1960)
KABC: Kaufman Assessment Battery for Children (Kaufman & Kaufman, 1983)

Table 2
WISC-R PIQ Mean Score Comparisons With Other Measures

Study	Subjects	WISC-R $\bar{x}$ (SD)	Criterion $\bar{x}$ (SD)
Moores, Weiss & Goodwin (1976)	60 preschool children 3yr interval LIPS first	 110.10	LIPS 116.60 range 82-157
Quigley (1969)	32 children $\bar{x}$ age=3.8 yr. 4yr interval LIPS first	 102.00	LIPS 114.00
Ritter (1976)	31 children $\bar{x}$ age=6.9 yr. HL=mild to moderate	 82.20 (16.00)	AALIPS 88.80 (18.10)
Hirshoren, Hurley & Hunt (1977)	59 children $\bar{x}$ age=10.6 yr. HL=prelingually deaf	 88.07 (17.84) range 52-129	H-NTLA 89.86 (16.53) range 57-131
Anderson & Sisco (1977)	1228 national sample age 6-0 to 16-11 HL=70dB+	 95.70 (17.55) n=1202	LIPS 97.44 (16.90) n=320
			H-NTLA 96.78 (16.64) n=219
James (1984)	34 children age 6-11 HL $\bar{x}$=85dB	 92.36	CPM 97.54
	50 children age 11-16.5 yr. HL $\bar{x}$=80dB	 101.41	SPM 99.06
Watson & Goldgar (1985)	71 children $\bar{x}$ age=13.20 yr. HL 44 prfound HL 22 severe HL 5 mod/severe	 99.13 (17.42)	H-NTLA 100.21 (22.36)

Table 2 continued

Study	Subjects	WISC-R $\bar{x}$ (SD)	Criterion $\bar{x}$ (SD)
Ullissi & Gibbons (1984)	40 children		LIPS
	$\bar{x}$ age=8.0 yr.	96.30	97.40
	HL=prelingually deaf 70dB+	(13.00)	(23.30)
Phelps & Ensor (1986)	49 children		H-NTLA
	$\bar{x}$ age=11.51 yr.	91.90	90.41
	HL=prelingually deaf	(16.93)	(20.46)
	(moderate to profound)	58-130 (deaf norms)	57-139
Porter & Kirby (1986)	49 children		KABC
	ages 7-12 yr.	107.92	98.8
	HL=prelingually deaf	n=25	(11.80)
	HL $\bar{x}$=102dB	pantomine/ gesture	
		108.38	96.8
		n=24	(10.90)
		(ASL)	
Watson *et al.* (1986)	53 children		H-NTLA
	$\bar{x}$ age=13.90 yr.	105.00	105.00
	HL=mod-severe to profound	(11.82)	(15.80)
Brooks & Riggs (1980)	40 children		WISC
	$\bar{x}$ age=10.90 yr.	90.35	96.08
	HL=40-80+dB	(19.81)	(17.57)

AALIPS: Arthur Adaptation of the Leiter International Performance Scale (Arthur, 1955)
CPM: Raven's Colored Progressive Matrices (Raven, 1965)
H-NTLA: Hiskey-Nebraska Test of Learning Aptitude (Hiskey, 1966)
LIPS: Leiter International Performance Scale (Leiter, 1948)
SPM: Standard Progressive Matrices (Raven, 1960)
KABC: Kaufman Assessment Battery for Children (Kaufman & Kaufman, 1983)

C. MODIFICATIONS IN TEST INSTRUCTIONS

The method of test administration has been determined to be a crucial variable in the psycho-educational assessment of the hearing impaired. Considerable work has been done in this area for it has been noted that misinterpretation can occur both with regard to the task demand given by the examiner and the task response given by the child. There has been some research providing empirical evidence that different procedures have a significant effect on test results. Graham and Shapiro (1953) compared the WISC Performance Scale results using two different administration methods. Two groups of children with normal hearing were individually matched with members of group I, the hearing impaired children, using the Goodenough Draw-A-Man Test. The WISC Performance Scale was administered individually to the members of Groups I and II through pantomime instructions. The members of Group III were given the WISC in accordance with the standardized instructions. Reported Performance IQs were:

Group I	Group II	Group III
96.1	95.5	101.5

The difference between groups II and III was found to be significant at the 5% level. The authors concluded that, because the three groups were originally equaled in intelligence as measured by the Goodenough, it would seem that the difference in WISC performance resulted from the method of administration.

Sullivan (1982) has also demonstrated that different administration modifications result in significantly different test scores for different categories of deaf children. From a school using Total Communication (TC) instructional methodology, 45 children with severe to profound prelingual hearing losses were sorted into three etiological groups. These were: a genetic group, a multiply handicapped group, and a group of unknown etiology. Three administration conditions were employed with the Performance Scale of the WISC-R: pantomime, visual aids, and total communication. Each of the 15 subjects in the three etiological groups was randomly assigned to one of three administration groups. General results indicated that the total communication administrations produced significantly higher scores. The implication drawn was that for hearing impaired children in residential and public schools who are instructed in a total communication methodology, the use of TC in the administration of the standardized subtests of the WISC-R Performance Scale would yield higher scores.

Bragman (1982) has presented a review of modified test instructions used with hearing impaired populations. The three methods customarily used have been pantomime; demonstration; and some form of language communication involving speech, lipreading, written words, or manual communication. In the past researchers have used these procedures singly or in combination. For instance Murphy (1957) used all three methods to some degree. He stated that speech should be used throughout the test so that students who could lipread would take advantage of it. He also employed some pointing and gesturing as well as additional demonstrations. Neuhaus (1967) believed only pantomime should be used regardless of how well the child was able to speechread. He felt that any advantage a good speechreader had over a poor one should be eliminated. Many researchers in the past have been rather vague in referring to procedures of test administration. Sattler (1988; 1982) has published two modified instruction procedures: the Pantomime instructions from Sullivan (1982) and a combination of Murphy (1957), Neuhaus (1967), and Reed (1970). Sattler (1988) has further recommended the use of TC to administer tests to children who use TC as their educational modality. Ray (1979) published an adaptation of the WISC-R Performance Scale administration procedures. These involve the use of supplemental items for demonstrations as well as Alternate Instructions for signing.

Each of the three categories of modified test instructions has certain advantages and disadvantages. The use of pantomime does not always clearly convey what is needed, nor is it necessarily free of language biases (Goetzinger and Houchins, 1969; Reed, 1970; Tweney and Hoemann, 1976). It has also been found to depress scores. (Graham and Shapiro, 1953; Sullivan, 1982). Pantomime administration would have considerable advantages if the test had been standardized in that manner. Otherwise, the effects of pantomime may be as variable as other test modifications.

The use of demonstration on extra practice items to modify test administration may have the serious effect of rehearsal. The effects on the validity of the test are not known. However, since the value of each test item depends to a great extent on the novelty of the task, it can be assumed that extra demonstration is very likely to alter the nature of the task. This type of test modification is seen as a more drastic alteration than pantomined instructions and less likely to produce valid results.

The use of language communication has the advantage of more closely following standardized procedures. The disadvantages are the variable language levels of the hearing impaired children. Unlike hearing children, hearing impaired children cannot be assumed to have a common language base.

There appears to be little consistency in the modifications used in actual assessment situations. Levine's (1974) survey of psychologists working with the hearing impaired revealed that 52% used mainly speech and 31% used mainly signs in test administration. The Anderson and Sisco (1977) norming study reported great inconsistency in mode of test instructions. Total Communication was used by 77.2% of the sample; 2.2% used Oral; 4.4% used fingerspelling with speech; 1.3% used only gestures; 7.3% used pantomime; 6.6% employed other, unidentified procedures.

In another survey McQuaid and Alovisetti (1981) reported that 71% of the respondents indicated they used Total Communication for test administration. It is important to note here that the use of TC does not necessarily mean the administrations were standardized. Total Communication is a style of communication encompassing many different communication modes. Audition, speechreading, signing and speaking, gestures, and pantomime are all incorporated. The exact proportions of each modality used in any communication event are not stipulated. Gesturing and pantomime would logically reflect the personal characteristics of the communicator. Some of the signs used may well be unique to the geographical region or even the specific school attended by the child. The results of a recent survey taken in the United States and Canada by Allen, Abraham and Stoker (1988) indicated that 48.8% of the respondents used both speech and sign simultaneously in assessment situations.

These research results appear to indicate a growing trend in the field of assessment of the hearing impaired: to use a Total Communication mode of communication. This may reflect the growing acceptance and use of TC in the education of the hearing impaired. As more children are taught using TC methodology, the more it becomes the most appropriate mode of communication in which to assess those children. In the United States, P.L.94-142, has stimulated additional incentives for the use of sign language. Section 12la 532 states:

> tests and other evaluation materials are provided and administered in the child's native language or other mode of communication, unless it is clearly not feasible to do so. (Federal Register, Aug 23, 1977 p. 42494).

It is this consideration of the individual abilities of the child which appears to be the most important determinant in the choice of administration mode. For the orally educated child, spoken communication should be employed. For children who use TC in school or who know sign language, a TC administration is considered appropriate (Sullivan and Vernon, 1979; Sattler, 1988).

The choice of administration mode becomes a judgement on the part of the examiner. What is most imperative is to choose the method most likely to decrease the incidence of misinterpretation between examiner and child. For this study, a form of Total Communication using Pidgin Sign English was chosen to be the administrative mode used for both tests. Stokoe (1970) found that there were different varieties of sign language within the deaf community. Stokoe described the language varieties in terms of a continuum ranging from American Sign Language (ASL) to English. Woodward (1973) developed this concept and used the term Pidgin Sign English (PSE) to describe the in-between varieties along the ASL-English continuum. He felt PSE was used as a bridge between the two languages. As such, the term PSE is used to describe signed communication that consists of some grammar from ASL and some grammar from English. There is some empirical evidence to suggest that PSE is used by deaf students in the classroom (Woodward 1973; Livingston, 1983). PSE could well be the communication form used by many parents and educators. Not all parents and educators are equally proficient in the use of ASL or Signed English. It is quite likely that there is a continuum of signing skills occurring with most communication being in the PSE form.

The Signing Consultant employed at the school from which the sample population of this study was drawn felt that a PSE format would be the most familiar form of communication for these children. Therefore, a careful PSE procedure was developed that closely followed the standardized administration procedures of each intelligence measure. For each of these tests, the procedures followed were identical for each child. In the judgement of the Sign Consultant, the PSE test administration procedures of the WISC-R Performance Scale and the SB:FE were consistent in their location along the ASL — English sign continuum.

CHAPTER III. METHODOLOGY

This study proposed to evaluate the appropriateness of selected subtests of the new Stanford-Binet Intelligence Scale: Fourth Edition for use with hearing impaired children whose educational methodology is Total Communication. The evaluation included correlations with the WISC-R Performance Scale and the academic measures of the SAT-HI.

Chapter III outlines the procedures of this study, describes the sampling and instrumentation used, and briefly describes the methods for data analysis.

A. PROCEDURE

For each subject who had parental permission to be included in this study, audiological information, medical information, and results of the previous year's SAT-HI academic measures were collected from the school files.

Each subject was administered both the selected subtests of the SB:FE and the WISC-R Performance Scale. Subjects were randomly assigned to one of two groups, and the order of test administration was counterbalanced for the two groups. Because of the effects of fatigue, the time of day was consistent for each child for both test administrations. The time interval between test administrations for any child was between two and three weeks. Testing was completed between April and June, 1988.

Administration Procedures: The standardized administration procedures were followed with modifications made in the communication mode. The communication mode chosen for test administration was a form of Total Communication combining the use of sign, fingerspelling and speech. The sign system used was a form of Pidgin Sign English. The signed administration format was carefully developed with the assistance of the Signing Consultant employed by the school. Local signs were incorporated where appropriate and the formats of both tests were judged to be consistent by the Consultant. Test administration was conducted by the researcher, a graduate student in Educational Psychology and Special Education trained in the use of both assessment instruments. In addition, this researcher holds a Diploma in the Education of the Hearing Impaired and has been a classroom teacher of hearing impaired children.

B. SAMPLING

Students between the ages of 8.0 and 16.11 years enrolled at a Provincial School for the Deaf had been invited to participate. The 38 subjects were a volunteer sample for whom parental permission had been obtained. They were enrolled in a residential/day School for the Deaf which uses Total Communication as the educational methodology. The subjects had severe to profound sensorineural hearing losses. There were 23 boys and 15 girls with an age range of 8.0 to 16.11 years included in the study. Medical records ascertained that the subjects had neither severe motor disabilities nor visual impairments. Etiologies were as follows:

Etiology	n
unknown	19
heredity	14
rubella	4
menningitis	1

This sample had a high proportion of children whose deafness can be attributed to causes of heredity.

With the exception of one student, all the subjects were considered prelingually deaf with onset before 2 years of age. The 1 postlingually deaf student's age of onset was 4 1/2 years. Analysis of this individual's data indicated no observable differences of his scores from the sample group mean scores.

Audiological information gathered from the medical records yielded the following sample description:

Better Ear Average (BEA) Mean	100dB
Standard Deviation	9.4dB
Range of Hearing Loss	75dB to 113dB

No student in this study had a prior WISC-R PIQ administration with the exception of a single student, who had undergone testing two years prior to this data collection.

C. INSTRUMENTATION

1. WISC-R Performance Scale

There are five subtests in this scale plus one supplementary subtest. Scaled scores with a mean of 10 and a standard deviation of 3 are obtained on each subtest. Together, these subtests form the Performance IQ which has a mean of 100 and a standard deviation of 15. For this study the 1974 Weschler norms were used.

Subtests

Picture Completion

This subtest consists of drawings of common objects. Each drawing lacks a single important feature. The drawings are presented one at a time and the child is required to point to the missing feature within a 20-second time limit. There are 26 drawings and the items are scored pass or fail. The subtest is discontinued after four consecutive failures.

Picture Arrangement

In this subtest the child is required to rearrange a set of pictures into a story sequence. There are 12 sets of the picture stories. Time limits increase progressively through the sets and fast responses receive bonus points. The subtest is discontinued after three consecutive failures.

Block Design

Using a set of patterned blocks, the child is required to duplicate a geometric design displayed on a card. There are 11 items of increasing difficulty. Time limits increase progressively and fast responses receive bonus points. The subtest is discontinued after two consecutive failures.

Object Assembly

For this subtest the child is required to assemble jigsaw puzzle pieces to form a common object. Each item has a time limit and fast responses receive bonus points. If the puzzle is incomplete after the time limit has expired, partial credit may be awarded. All children receive all 4 puzzle items.

Coding

This is a paper and pencil task. A Code Key is displayed at the top of the page, and the child is required to fill in as many code items as possible in the blank boxes below as possible within two minutes.

Mazes

This is a supplementary test and is usually not included in the composite score. In this paper and pencil task, the child is required to accurately trace a path from the middle of a maze out to the exit. All items are timed and the number of errors made determines the child's score.

2. Stanford-Binet Intelligence Scale: Fourth Edition–Selected Subtests

The subtests selected for this study were considered to be nonverbal performance type tests. The Verbal Reasoning Area subtests were excluded because in this initial investigation of the SB:FE, their inclusion to form an estimate of intelligence would not be supported (Vernon and Brown, 1964; Sullivan and Vernon, 1979). Neither was the Quantitative subtest included in this study because of the language content of the word problems as well as the reading requirement. It is well supported in the literature that reading levels in the hearing impaired population are significantly lower than in the non-exceptional population (King and Quigley, 1985).

Order of Administration

The order of administration followed the sequence in the test response booklet. For this study, in order to gather as much data as possible, it was decided that each child would attempt each subtest. This is a departure from the organization of the SB:FE. Normally, the Vocabulary subtest is used as a routing test to determine, together with age, which subtests will be included in the test battery. The Vocabulary subtest also determines the entry level into these subtests. The entry levels used in this study will be noted in the subtest descriptions listed below. The instructions for the adaptive testing procedures described in the administration manual were followed in this study. If the basal level (four consecutive correct responses) could not be achieved at the entry level, then testing would progress backwards until a basal level was achieved. For all subtests, test administration was discontinued when three out of four items were failed in two consecutive levels, as in standardized instruction.

Selected Subtests

Abstract/Visual Reasoning Area

Pattern Analysis

Using a set of patterned blocks, the child is required to duplicate a geometric pattern which is initially modelled by the examiner and on later items is displayed on a card. The pattern must be correctly completed within a certain time limit. Entry level was determined by the basal level (four consecutive correct responses) attained in the preceding Bead Memory subtest.

Copying

In this subtest the child is required to reproduce geometric designs. Depending on entry level, the nature of the task varies. Initially, young children are required to reproduce with unpatterned blocks, a model demonstrated by the examiner. The task then changes to a paper and pencil task where the child copies a design from a displayed line drawing. Entry level for each child was the first item of the paper and pencil task.

Matrices

Each item in this subtest consists of a configuration of cells containing three figures and one blank cell. The child selects the best alternative to complete the matrix. In each item, the missing figure for the blank cell has some logical relationship to the other figures. For all children entry level was the first item.

Paper Folding and Cutting

For each item, the child first looks at a sequence of drawings showing the stages of a piece of paper as it is folded and cut. The child then selects the best alternative to represent how the folded and cut paper would look unfolded. For all children entry level was the first item.

Quantitative Reasoning Area

Number Series

The child is shown a sequence of numbers arranged according to a certain rule, and the child must indicate which two numbers would come next in the series. For each child entry level was the first item.

Equation Building

Each item consists of a series of numbers followed by operation signs. The numbers and signs must be rearranged to make a true number sentence. For each child entry level was the first item.

Short-Term Memory Area

Bead Memory

Using beads of various shapes and colours, the child is required to reproduce from memory, a picture of beads arranged on a stick. The picture is displayed for five seconds. For each child entry level was initially item 11, which is the first item using the beads on a stick format.

Memory for Objects

The child views pictures of objects presented in a specific order at one second intervals. Then from a picture containing those objects plus several other distracting objects, the child must identify those objects displayed in the correct sequence. For each child entry level was the first item.

3. Stanford Achievement Test (7th Ed.), Special Edition for Hearing Impaired Students (Gardner *et al.*, 1982)

The SAT-HI is a norm-referenced achievement test. Separate norms have been provided for Hearing Impaired students. Entry into different tests is determined by grade levels. Tests recommended for use with hearing impaired students are:

Reading Comprehension

A sentence completion test with a multiple choice format.

Spelling

For each item, a group of phrases are presented, each one having one underlined word. Students are to identify the incorrectly spelled underlined word.

Language

For each item, a group of words are presented. The student has four options: identify the group as (1) a complete sentence, (2) a run-on sentence, or (3) and (4) which are alternate choices to complete the sentence.

Word Reading

For each item, students match a picture with the correct word.

Concepts of Number

Numerical concept questions are provided with multiple-choice answer format.

Math Computations

Computation questions are presented with a multiple choice answer format.

Math Applications

Word problems are presented with a multiple choice answer format.

D. ANALYSIS

All analyses for the SB:FE were computed using all subtests for each subject. For younger children who obtained a score in subtests beyond their norms, the Standard Age Score was obtained from the next highest age norm table available. For comparison purposes, each subject's test was rescored using an alternate method with the Bead Memory basal level determining the entry level for the remaining tests, as suggested in the SB:FE Examiner's Handbook (Delaney & Hopkins, 1987), for testing examinees who have limited English proficiency or are non-language proficient. This comparison is displayed in Table B1 of Appendix B. For both test instruments, the published norm tables were used to obtain the Scaled Scores or Standard Age Scores and their composites.

Correlations between the selected subtests of the SB:FE, the WISC-R Performance Scale and the academic measures of the SAT-HI were calculated using the SPSS-X Data Analysis System, (SPSS Inc., 1988). The level of significance was set at $p=0.05$ for this study. An exception was made for the calculations involving the Language measure of the SAT-HI which had less than 10 subjects. Therefore, to minimize the risk of inflating Type II error to untenable levels, alpha was relaxed to $p=0.10$.

The t-test for significance was calculated using the SPSS-X computer program to determine systematic differences between the selected subtests of the SB:FE and the WISC-R Performance Scale.

The patterns of subtest intercorrelations for the selected subtests of the SB:FE and the WISC-R Performance Scale were compared using the computer program MULTICORR (Steiger, 1979).

Item analysis on the selected subtests of the SB:FE was performed using the computer program LERTAP (Galan and Nelson, 1986).

CHAPTER IV. RESULTS

Chapter IV will present the results of the study and describe the results in relationship to the hypotheses tested.

Hypothesis 1 stated that for the hearing impaired sample in this study, the pattern of correlations between the WISC-R Performance Scale I.Q. (PIQ) and the Abstract/Visual, Quantitative and Short-Term Memory Area scores of the Stanford-Binet: Fourth Edition will not be significantly different from that reported for the non-exceptional sample.

In order to test this hypothesis, Pearson R coefficients were obtained for the hearing impaired group. Table 3 displays these correlation coefficients and also the data reported for the non-exceptional sample. To determine whether the pattern of correlations for the hearing impaired group was significantly different from the pattern reported for the non-exceptional sample, a test for the null hypothesis regarding the difference between independent correlation coefficients was employed. This test was carried out according to procedures described by Glass and Hopkins (1984, p. 307).

The difference in correlation coefficients for the hearing impaired group and the non-exceptional sample is significant for the SB:FE Abstract/Visual Area with the PIQ. The correlation of the two test instruments for the hearing impaired group is significantly higher than the correlation reported for the non-exceptional sample. The correlation between the Quantitative and the PIQ, and the Short-Term Memory and PIQ for the hearing impaired group was computed to be not significantly different from the non-exceptional data. Because the pattern of correlations for the hearing impaired group was different from that reported for the non-exceptional sample, Hypothesis 1 was rejected.

Table 3

Correlations Between Area Scores of the Stanford-Binet: Fourth Edition and the WISC-R PIQ

	Stanford-Binet Intelligence Scale (Fourth Edition)		
	Abstract/Visual r	Quantitative r	Short-Term Memory r
Hearing Impaired WISC-R PIQ	**82** n=38	74 n=35	72 n=38
Non-exceptional WISC-R PIQ	67	63	63

Note: All entries rounded to 2 significant figures, decimals omitted. All coefficients are significantly different from zero at α=.05. However, when comparing the hearing impaired coefficents to the non-exceptional coefficients, only the coefficients for the Abstract/Visual Scale (82 versus 67) are significantly different at α=.05.

Hypothesis 2 stated that for the hearing impaired group, the pattern of subtest intercorrelations of the SB:FE will not be significantly different from that reported for the non-exceptional sample.

Hypothesis 3 stated that for the hearing impaired group in this study, the pattern of subtest intercorrelations of the WISC-R PIQ will not be significantly different from that reported for the non-exceptional sample.

To test these hypotheses, the computer program MULTICORR (Steiger, 1979) was used. Results of this analysis are summarized in Table 4. These results indicate that the pattern of intercorrelations of the SB:FE subtests for the hearing impaired sample is not significantly different than for the non-exceptional sample. Hypothesis 2 can therefore be accepted. For the WISC-R Performance Scale subtests, the pattern of intercorrelations for the hearing impaired sample is not significantly different than for the non-exceptional sample. Hypothesis 3 can therefore be accepted.

Table 4

Subtest Pattern Intercorrelations Between the Hearing Impaired Sample and the Non-exceptional Sample

Hypothesis	Chi Square	df	p
2 (SB:FE)	18.41	21	.62
3 (WISC-R PIQ)	21.17	15	.13

Hearing Impaired Sample: n=38

Exploratory research questions were addressed as part of this study.

1. To investigate item characteristics of the various subtests of both instruments, item analysis was conducted using the computer program LERTAP (Galan and Nelson, 1986). The exploratory item analysis provided the results displayed in Table 5. For a more comprehensive display of the item analysis of each subtest, the reader is referred to Appendix A. Examination of the subtest analysis revealed that the only remarkable finding occurred in the Equation Building subtest. All subjects failed the first item in this subtest. As the SB:FE has been constructed to reflect a stepwise progression of item difficulty, the failure of the first item is unexpected.

Table 5

Item Analysis Characteristics for the SB:FE

Subtest	Hoyt Estimate	Raw Score Mean[1]	Standard Deviation	Standard Error of Measure
Pattern Analysis	.92	31.92	6.44	1.84
Copying	.78	20.21	3.37	1.56
Matrices	.93	8.68	5.93	1.56
Paper Folding & Cutting n=29	.91	2.76	3.59	1.07
Number Series n=35	.90	8.79	4.84	1.53
Equation Building n=15	.85	1.16	2.14	0.81
Bead Memory	.86	21.53	5.03	1.84
Memory for Objects	.73	5.97	1.99	1.00

[1]No comparable values are given in manual; only transform score means provided.

Note: For all subtests: Cronbachs Alpha for Composite = .88
n=38 unless otherwise indicated.

2. For the hearing impaired group in this study, will either instrument systematically yield greater Standard Age Scores? To conduct this analysis, the WISC-R Performance Scale IQ was converted to the SB:FE scale using the following formula:

$$Z = [\ ((\text{WISC-R PIQ} - 100)/15) \times 16\ ] + 100$$

A correlated groups t-test was performed between the converted WISC-R PIQ scores and each Standard Age Score of the SB:FE. Table 6 displays the results of this analysis. The WISC-R Performance Scale IQ scores are almost a

Table 6

Comparison of Standard Age Area Scores of the SB:FE and Converted WISC-R PIQ Scores for the Hearing Impaired

Test	n	Mean	Standard Deviation	Mean (Difference)	t	p
WISC-R PIQ	38	99.89	21.27	14.57	7.27	<0.01
SB:FE Partial Composite		85.32	17.58			
WISC-R PIQ	38	99.89	21.27	12.55	6.27	<0.01
SB:FE Abstract/Visual		87.34	17.49			
WISC-R PIQ	35	100.21	21.16	11.24	4.46	<0.01
SB:FE Quantitative		89.97	14.80			
WISC-R PIQ	38	99.89	21.27	15.42	6.26	<0.01
SB:FE Short-Term Memory		84.47	18.56			

$a=.05$

standard deviation higher than any of the SB:FE Area Scores. It can be stated that there is a significant systematic difference between these two instruments for this hearing impaired group. The alternate scoring method for the SB:FE, displayed in Table B1, produced a group partial composite mean of 85.24. The different scoring methods did not yield significantly different group means.

3. The pattern of subtest intercorrelations between the WISC-R Performance Scale and the SB:FE was investigated. The computer program LERTAP (Galan and Nelson, 1986) was used in this data analysis. The results of this investigation are displayed in Table 7.

The median intercorrelations of the various subtests of the WISC-R and the SB:FE produced coefficients ranging from Picture Completion (WISC-R) at .64

Table 7

Subtest and Area Intercorrelations Between the SB:FE and the WISC-R PIQ

WISC-R PIQ	SB:FE												
	PT	CP	MX	PF	AV	NS	EB	QA	BM	MO	ST	PO	SM
PC	**64**	**63**	**70**	**58**	**78**	**60**	**62**	**65**	**65**	**65**	**62**	**74**	64
PA	**56**	**58**	**67**	**32**	**66**	**68**	**79**	**72**	**64**	**54**	**67**	**73**	61
BD	**78**	**75**	**55**	**57**	**79**	**70**	**54**	**71**	**68**	**49**	**64**	**76**	63
OA	**55**	**65**	**49**	**38**	**63**	**57**	**54**	**59**	**56**	**30**	**50**	**60**	55
CD	**46**	**38**	**41**	**35**	**50**	**42**	39	**45**	**49**	**46**	**59**	**55**	42
MZ	24	23	10	12	21	**31**	-5	**32**	22	26	27	**29**	33
PS	**73**	**73**	**68**	**53**	**81**	**71**	**74**	**74**	**72**	**53**	**72**	**81**	
SM	56	61	52	37		59	54		60	48			

Note: All entries rounded to two significant figures, decimals omitted, boldface coefficients significant at $\alpha=.05$
n=38 unless otherwise indicated below

PC: Picture Completion
PA: Picture Arrangement
BD: Block Design
OA: Object Assembly
CD: Coding
MZ: Mazes
PS: Performance Scale I.Q.
SM: Subtest Median
PT: Pattern Analysis
CP: Copying
MX: Matrices
PF: Paper Folding & Cutting (n=29)
AV: Abstract/Visual Area
NS: Number Series (n=35)
EB: Equation Building (n=15)
QA: Quantitative Area (n=35)
BM: Bead Memory
MO: Memory for Objects
ST: Short-Term Memory Area
PO: Partial Composite

to Mazes (WISC-R) at .33. The following list displays in rank order the six highest median subtest intercorrelations:

Subtest	r
Picture Completion (WISC-R)	.64
Block Design (WISC-R)	.63
Picture Arrangement (WISC-R)	.61
Copying (WISC-R)	.61
Bead Memory (SB:FE)	.60
Number Series (SB:FE)	.59

The highest single subtest correlational coefficients were as follows: Block Design (WISC-R) correlated .79 with the Abstract/Visual Area (SB:FE) and .78 with Pattern Analysis (SB:FE). Picture Arrangement (WISC-R) correlated .79 with Equation Building (SB:FE). Picture Completion correlated .78 with the Abstract/Visual Area score of the SB:FE.

The highest correlation coefficients of .81 were computed between both the WISC-R Performance IQ and the Abstract/Visual Area and between the WISC-R Performance Scale IQ and the Partial Composite of the SB:FE.

4. Both the WISC-R Performance Scale and the SB:FE were correlated with academic measures from the Stanford Achievement Test – Hearing Impaired (SAT-HI). For this analysis the Pearson R correlational coefficient was computed for each measure.

Of the various WISC-R Performance Scale subtests, Picture Arrangement produced the most and the highest significant correlations. Object Assembly had no significant correlations with any of the academic measures.

Of the SAT-HI measures, Number Concepts and Reading produced more significant correlations (see Table 8). Language and Word Recognition produced no significant correlations with any of the WISC-R Performance Scale results.

Table 8

Correlations of the WISC-R PIQ and the SAT-HI

	SAT-HI						
WISC-R PS	Reading Comp	Spelling	Language	Word Reading	Number Concepts	Math Comp	Math Applic
Picture Completion	**49** n=32 p=.002	**39** n=20 p=.043	28 n=7 p=.269	35 n=18 p=.075	**48** n=32 p=.003	27 n=32 p=.005	**54** n=27 p=.002
Picture Arrangement	**54** n=32 p=.001	**47** n=20 p=.019	29 n=7 p=.266	01 n=18 p=.488	**53** n=32 p=.001	**43** n=32 p=.007	**46** n=27 p=.007
Block Design	**38** n=32 p=.017	31 n=20 p=.094	-29 n=7 p=.266	21 n=18 p=.200	**32** n=32 p=.037	17 n=32 p=.171	**25** n=27 p=.104
Object Assembly	23 n=32 p=.105	31 n=20 p=.089	-35 n=7 p=.221	-05 n=18 p=.416	24 n=32 p=.096	20 n=32 p=.137	24 n=27 p=.119
Coding	**31** n=32 p=.042	23 n=20 p=.161	20 n=7 p=.334	35 n=18 p=.076	**30** n=32 p=.046	23 n=32 p=.099	**33** n=27 p=.046
Mazes	28 n=32 p=.063	**41** n=20 p=.035	-52 n=7 p=.117	07 n=18 p=.386	**37** n=32 p=.017	**37** n=32 p=.019	31 n=27 p=.056
Performance I.Q.	**46** n=32 p=.004	**42** n=20 p=.032	-01 n=7 p=.489	20 n=18 p=.219	**44** n=32 p=.005	**32** n=32 p=.037	**43** n=27 p=.012

Note: Boldfaced coefficients significant at α=.05, exrept where $n \leq 10$ then α=.10. All entries rounded to two significant figures, decimals omitted.

Table 9 displays the results of the correlations between the SB:FE and the SAT-HI. Among the SAT-HI measures, Reading Comprehension, Language and Number Concepts produced the greatest number of significant correlational coefficients. Language produced the highest coefficients. Of the various SB:FE subtests, the Short-Term Memory Area produced the most significant correlations. The Paper Folding and Cutting subtest yielded no significant correlations.

Table 9

Correlations of the SB:FE and the SAT-HI

	SAT-HI						
SB:FE	Reading Comp	Spelling	Language	Word Reading	Number Concepts	Math Comp	Math Applic
Pattern Analysis	**40** n=32 p=.011	21 n=20 p=.186	19 n=7 p=.341	24 n=18 p=.170	**40** n=32 p=.012	18 n=32 p=.158	29 n=27 p=.072
Copying	**32** n=32 p=.037	20 n=20 p=.197	-17 n=7 p=.360	21 n=18 p=.202	24 n=32 p=.096	07 n=32 p=.352	24 n=27 p=.113
Matrices	**49** n=32 p=.002	18 n=20 p=.229	52 n=7 p=.144	21 n=18 p=.204	**38** n=32 p=.017	27 n=32 p=.069	32 n=27 p=.053
Paper Folding & Cutting	10 n=26 p=.314	-07 n=17 p=.391	27 n=7 p=.283	47 n=12 p=.064	05 n=26 p=.413	-25 n=26 p=.107	-02 n=22 p=.463
Abstract/Visual Reasoning Area	**42** n=32 p=.008	22 n=20 p=.182	32 n=7 p=.241	31 n=18 p=.107	**37** n=32 p=.020	16 n=32 p=.193	31 n=27 p=.061
Number Series	**39** n=31 p=.016	24 n=19 p=.164	**63** n=7 p=.064	12 n=17 p=.326	29 n=31 p=.060	09 n=31 p=.312	27 n=26 p=.091

Table 9 continued

SB:FE	Reading Comp	Spelling	Language	Word Reading	Number Concepts	Math Comp	Math Applic
	SAT-HI						
Equation Building	33 n=14 p=.126	**83** n=8 p=.003	40 n=5 p=.254	-79 n=4 p=.105	36 n=14 p=.103	-03 n=14 p=.462	29 n=10 p=.212
Quantitative Reasoning Area	**41** n=31 p=.011	26 n=19 p=.138	**57** n=7 p=.093	16 n=17 p=.264	**36** n=31 p=.024	12 n=31 p=.253	31 n=26 p=.064
Bead Memory	50 n=32 p=.002	23 n=20 p=.162	**63** n=7 p=.066	15 n=18 p=.274	27 n=32 p=.067	14 n=32 p=.225	24 n=27 p=.110
Memory for Objects	43 n=32 p=.007	11 n=20 p=.323	**64** n=7 p=.059	01 n=18 p=.480	**33** n=32 p=.031	21 n=32 p=.119	22 n=27 p=.132
Short-Term Memory Area	**52** n=32 p=.001	21 n=20 p=.186	**65** n=7 p=.059	09 n=18 p=.355	**33** n=32 p=.031	19 n=32 p=.145	27 n=27 p=.085
Partial Composite	**48** n=32 p=.003	24 n=20 p=.155	**55** n=7 p=.100	21 n=18 p=.204	**37** n=32 p=.018	17 n=32 p=.176	31 n=27 p=.059

Note: Boldfaced coefficients significant at $\alpha=.05$, except where $n\leq 10$ then $\alpha=.10$. All entries rounded to two significant figures, decimals omitted.

CHAPTER V. SUMMARY AND DISCUSSION

In this final chapter, the purposes, procedures, and results of the study are summarized. The major findings and their implications are then discussed and an orientation for further research is suggested.

A. SUMMARY

The present study proposed to evaluate the appropriateness of selected subtests of the new Stanford-Binet Intelligence Scale – Fourth Edition (SB:FE) for use with hearing impaired children whose educational methodology is Total Communication. It was reasoned that if the test characteristics of the SB:FE were similar for both the hearing impaired group and the non-exceptional standardization sample then the test could be used for its intended purpose. In the field of education, tests of intelligence are traditionally used to estimate general cognitive ability and to predict potential levels of academic achievement.

The subjects used in this study were enrolled in a residential/day School for the Deaf whose educational methodology was Total Communication. The subjects were tested on both the selected nonverbal subtests of the SB:FE and the Performance Scale of the WISC-R. The order of test administration was counterbalanced and a consistent Total Communication format was used for both tests.

To examine test characteristics of the SB:FE, the data gathered from the group of hearing impaired children were compared with the standardization data for the non-exceptional sample as reported in the technical manual. These comparisons involved correlations with the Performance Scale of the WISC-R, patterns of subtest intercorrelations, as well as exploratory item analysis of the

SB:FE. Correlations were also computed between each cognitive measure and measures of academic achievement. In addition, the Area Scores of the SB:FE were compared to the Performance IQ of the WISC-R to detect any systematic differences between these two instruments.

B. DISCUSSION OF RESULTS

Hypothesis 1 stated that for the hearing impaired sample in this study, the pattern of correlations between the Performance Scale of the WISC-R and the Abstract/Visual, Quantitative, and the Short-Term Memory Area scores of the SB:FE would not be significantly different from that reported for the non-exceptional sample. It is recalled that for the non-exceptional sample, the correlations were reported as being .63 between the Short-Term Memory Area and the WISC-R PIQ, .63 between Quantitative and the PIQ, and .67 between Abstract/Visual and the PIQ. The pattern is that of similar correlations between the Short-Term Memory and the Quantitative Areas with the PIQ, with a relatively higher correlation between the Abstract/Visual Area and PIQ. The results indicate that this general pattern is mirrored in the hearing impaired sample of this study. The correlation between Short-Term Memory and the PIQ was .72, between Quantitative and the PIQ it was .74, and between Abstract/ Visual and the PIQ it was .82. Comparing the correlations of the hearing impaired sample and the non-exceptional sample yielded the result that there were no significant differences between either the Quantitative Reasoning Area and the WISC-R Performance IQ or between the Short-Term Memory Area and the WISC-R Performance IQ. In this sample however, the Abstract/Visual and PIQ correlation coefficient was great enough to be significantly higher than its counterpart in the non-exceptional sample.

This pattern for both the non-exceptional sample and the hearing impaired sample is logical considering the nature of the subtests from either instrument. The Performance Scale of the WISC-R does not have a separate subtest dealing with visual short-term memory as does the SB:FE. Nor does the Performance Scale have a specific subtest dealing with numerical concepts as does the Quantitative Area of the SB:FE. The Arithmetic subtest of the WISC-R is to be found in its Verbal Scale rather than in the Performance Scale. The Abstract/ Visual Reasoning Area of the SB:FE appears to have more content similarity with the WISC-R Performance Scale. For example, Block Design of the WISC-R and Pattern Analysis of the SB:FE both require the reproduction of geometric designs with patterned blocks. The subtest intercorrelations for these two intelligence measures produced a coefficient of .78 for the hearing impaired children of this sample, confirming the suggestion of apparent content similarity. Sattler (1988) has suggested that the Performance Scale of the WISC-R might be considered to be an index of nonverbal and fluid intelligence involving immediate problem solving ability. In most subtests the stimuli are presented visually and may be described as perceptual organization skills. This description coincides with the information provided in the Administrative Manual of the SB:FE that describes the three-level hierarchical model of the structure of cognitive abilities upon which the SB:FE has been based. According to this model, the Abstract/ Visual Reasoning subtests are from the second level ability termed Fluid/Analytic Abilities (Thorndike, Hagen and Sattler, 1986a). These abilities are thought to require cognitive skills necessary for solving new problems that involve figural or other nonverbal stimuli. It is apparent that the intent of these areas of both instruments is to assess the construct of fluid intelligence. Both tests also use predominantly nonverbal performance-type tasks using visual perceptual abilities to tap this dimension of intelligence. Considering these similarities of intent and means for the WISC-R Performance Scale and the Abstract/Visual Area of the SB:FE, the results indicating higher correlations for these areas would reasonably be expected.

Hypothesis 2 stated that for the hearing impaired sample in this study, the pattern of subtest intercorrelations among the Stanford-Binet: Fourth Edition will not be different from that reported for the standardization sample. Results of this study found that there were no significant differences found for the hearing impaired group. This suggests that the dimensions of intelligence being tapped by the SB:FE are operating in a similar manner for both populations. This in turn suggests the subtest specificity is similar for both populations. The hearing impaired children do not seem to be substituting any cognitive abilities or skills that are different from those used by the standardization sample.

The exploratory item analysis results confirm this general evidence of similar subtest specificity. This exploratory investigation revealed the progression of item difficulty and test reliability throughout each subtest was found to be satisfactory with the exception of the Equation Building subtest. A discussion of the possible reasons for this finding is deferred to a later analysis of the group subtest means profile.

The similarity of patterns of subtest intercorrelations for both populations, as well as the acceptable item characteristics of most of the subtests indicates thus far that this instrument operates in a similar manner for the hearing impaired sample as it does for the standardization sample.

Hypothesis 3 compared the subtest intercorrelations of the WISC-R Performance Scale with the standardization sample. The study results found no significant difference for the hearing impaired group. The dimensions of intelligence tapped by the WISC-R Performance Scale appear to be operating in a similar manner for both populations. It will be recalled that the choice of administration mode for the testing conducted in this study was based on the work of Sullivan (1982) who found increased mean scores and reduced subtest differences when using a TC communication mode with children whose educational methodology was TC. The results of this analysis and the group subtest profile displayed in Figure 1 seem to support Sullivan's findings. As well, factor analytic research has revealed similar "principal factors extracted from deaf and hearing samples" on the WISC-R (Braden, 1985, p. 378). It seems reasonable to conclude that the choice of administration mode was correct and that this particular sample of hearing impaired children can be considered to be consistently average.

The exploratory research questions compared these two cognitive measures in several ways. The means of the total test scores were compared for significant differences, the subtests of both measures were intercorrelated and, in addition, each cognitive measure was correlated with the SAT-HI academic measures. The results of these investigations have produced some questions regarding the appropriateness of the SB:FE with this population of hearing impaired children.

The most significant comparison was the total test score means comparison. When the test results of the hearing impaired study sample were converted to the same scale and tested for significance, the differences were consistently almost 15 points which was significant at the .001 level. The scores

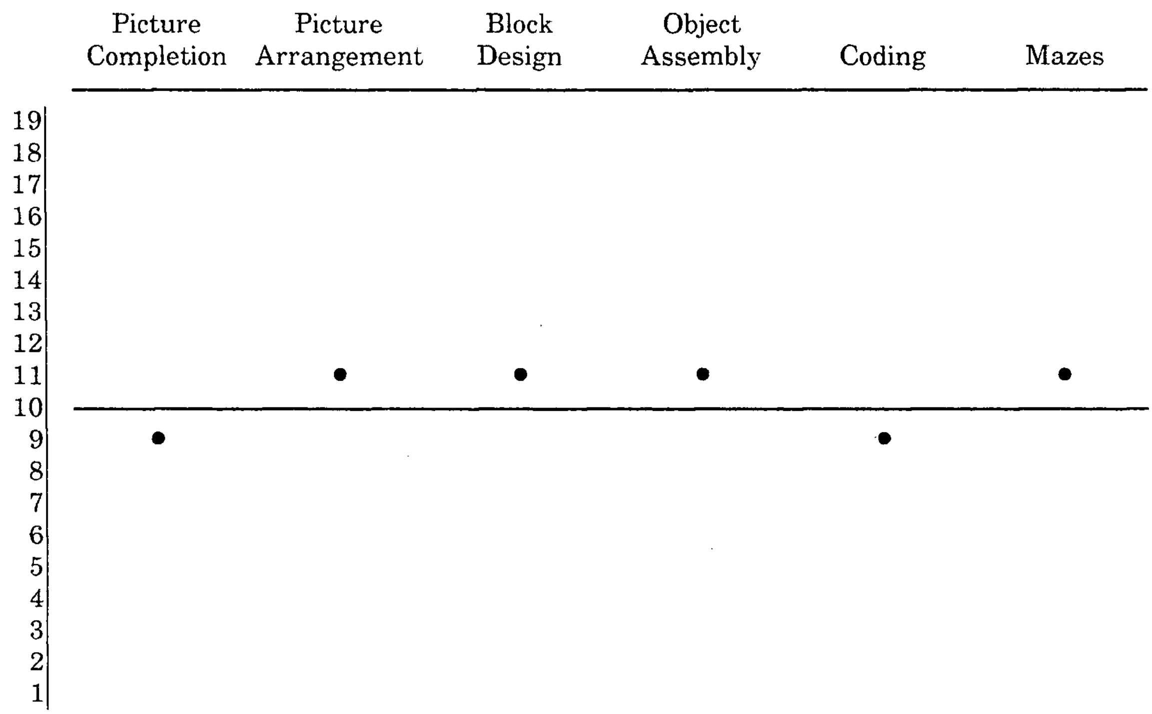

Figure 1

WISC-R PIQ Subtest Profile[1]

[1]Mean of 10, standard deviation of 3

from the Stanford-Binet were a standard deviation lower than the scores from the WISC-R Performance IQ. Such a significant difference has not been reported for the non-exceptional population. As a basis for comparison, it is noted in the Technical Manual of the Stanford-Binet that for a non-exceptional sample the mean WISC-R Performance IQ was 105.3 and from the SB:FE, the Abstract/ Visual Reasoning Area was 98.9, the Quantitative Area was 102.1, and Short-Term Memory was 102.6. Thus the score discrepancies found between these two instruments in this hearing impaired group are striking. A difference of this magnitude between two cognitive measures certainly precludes their use as interchangeable instruments. Estimates of cognitive ability are often an important factor in educational placement decisions. A child classified according to one test would most likely receive a different classification if judged by the other test. This would result in considerable confusion and lack of consistency in placement decisions for hearing impaired children. This is especially true because of the recommendation of plurality of assessment measures made by Sullivan and Vernon (1979), who advocated that one test of cognitive ability be used to confirm the results of another. This is deemed necessary in the hearing impaired populations because of the unique problems associated with testing the cognitive abilities of hearing impaired children.

The conclusion that these two instruments cannot be used interchangeably for this hearing impaired sample supports, in part, Sattler's (1988) conclusion regarding the interchangeability of the WISC-R and the SB:FE for the non-exceptional population. He cites the correlations as being .66 to .83 between the WISC-R Full Scale and the SB:FE Composite. Sattler concludes that the two tests yield scores that are approximately equal but they are not interchangeable. It is therefore concluded that the values of the correlations produced from this study are similarly not high enough to permit recommendation for their use together as interchangeable instruments of cognitive ability.

The marked mean score discrepancy between these instruments adds further support to this conclusion. It is again noted that these large mean score differences were not found in the correlational studies in the standardization sample. At this point in the discussion, the possible reasons for a discrepancy of this magnitude in the hearing impaired sample should be suggested. The question to be addressed then is, what are the differences between these two cognitive measures for the hearing impaired sample? It seems reasonable to suspect that the differences lie in the nature of the subtest task demands. It will be recalled that for this study, sources of external contamination were anticipated and steps were undertaken to minimize their effects. Subjects with identified motor or visual impairments were excluded from the study. It was also anticipated that the WISC-R PIQ results would reveal any unusually unique population characteristics for this study's hearing impaired sample. The administration procedures were also identified as being a possible source of contamination. It was decided that for this study, a consistent Total Communication format would be used for both test administrations. As explained in Chapter II, Sullivan's (1982) work was influential in this decision. Sullivan had found increased scores and minimized subtest differences when a Total Communication format was used in the test administration of the WISC-R PIQ with hearing impaired children using Total Communication as the educational methodology. The results of this study concur with Sullivan's findings. The group subtest profile displayed in Figure 1 shows this group to be average in all subtests compared to the standardization sample. This finding, in addition to supporting the Total Communication format issue, also provides information about this group's unique characteristics. The results of the WISC-R PIQ administration have revealed this particular group of hearing impaired children to be consistently average in the measured abilities. The mean Performance IQ for this group was close to the mean Performance IQ for the norming sample as reported in the manual. The standard deviation was considerably greater for this hearing impaired group. This is consistent with the findings from other studies using the WISC-R PIQ which found greater standard deviations for their hearing impaired samples (Ritter, 1976; Hirshoren *et al.*, 1977; Anderson & Sisco, 1977; Brooks & Riggs, 1980; Watson & Goldgar, 1985; Phelps & Ensor, 1986). These larger standard deviations values may be reflecting the greater possibility of additional handicaps due to the different etiologies responsible for the hearing impaired. The standard deviation found in this study was even larger than those previously reported, which may also be a characteristic of this particular sample. It can be concluded that in the cognitive abilities tested by the WISC-R PIQ, the hearing impaired group in this study was similar to their non-exceptional age peers. This finding supports previous research reviewed in Chapter II that suggests hearing impaired children are in the average range on measures of nonverbal intelligence (Vernon 1968).

For the purposes of this study it can be concluded that this particular group of hearing impaired children appear to be average in their cognitive abilities as measured by the WISC-R PIQ. It can further be suggested that the Total Communication format appeared to be appropriate. Since the TC formats for both instruments were judged as being consistent by the Sign Consultant, It seems justifiable to assume that the TC format for the SB:FE was similarily appropriate.

The selected subtests of the SB:FE may be measuring a dimension of cognitive ability that this hearing impaired sample are relatively low in when compared to their age appropriate peers in the standardization sample. Referring to Table 6 it will be remembered that the mean score differences are fairly consistent across all areas. For the hearing impaired sample the scores were: 87 in Abstract/Visual reasoning, 89 in Quantitative reasoning, and 85 in Short-Term Memory. This uniformity of score results seems to indicate there may be an underlying characteristic present in these selected subtests of the SB:FE that is responsible for the generally depressed scores. It is also noted that the standard deviation was reduced as well for this hearing impaired sample in comparison to the WISC-R PIQ results.

The group mean subtest profile displayed in Figure 2 permits a more detailed picture of the results. The two lowest scoring subtests; Equation Building and Copying, may have had interfering external factors operating. In the Equation Building subtest, item analysis revealed all subjects in the study group failed the first item. This is interpreted as reflecting a difference in curriculum content for this particular group of hearing impaired children. This subtest contains items consisting of a series of numbers and symbols that are to be rearranged to make a true number sentence. The first item required an arrangement which needed operations on both sides of the equal sign. The number sentence $5+2=3+4$ is an example of this arrangement. The example items for this subtest did not follow this format, but rather had only one digit following the equal sign. It could be that the hearing impaired children have been so accustomed to performing mathematical operations in the format demonstrated in the example items that an alternate format did not appear as a possibility. The Copying subtest was also observed to reflect a possible difference in curriculum emphasis for the hearing impaired children. In the administration of this subtest, the examiner noted a consistent attitude of disregard for precise duplication of the geometric designs. The scoring criteria of this subtest are very strict and designs may be failed, for example, for incomplete closure of a circle, or for a tilt in excess of 10 degrees of a design on the page. As a group, the hearing impaired students seemed inattentive to these fine details. A possible

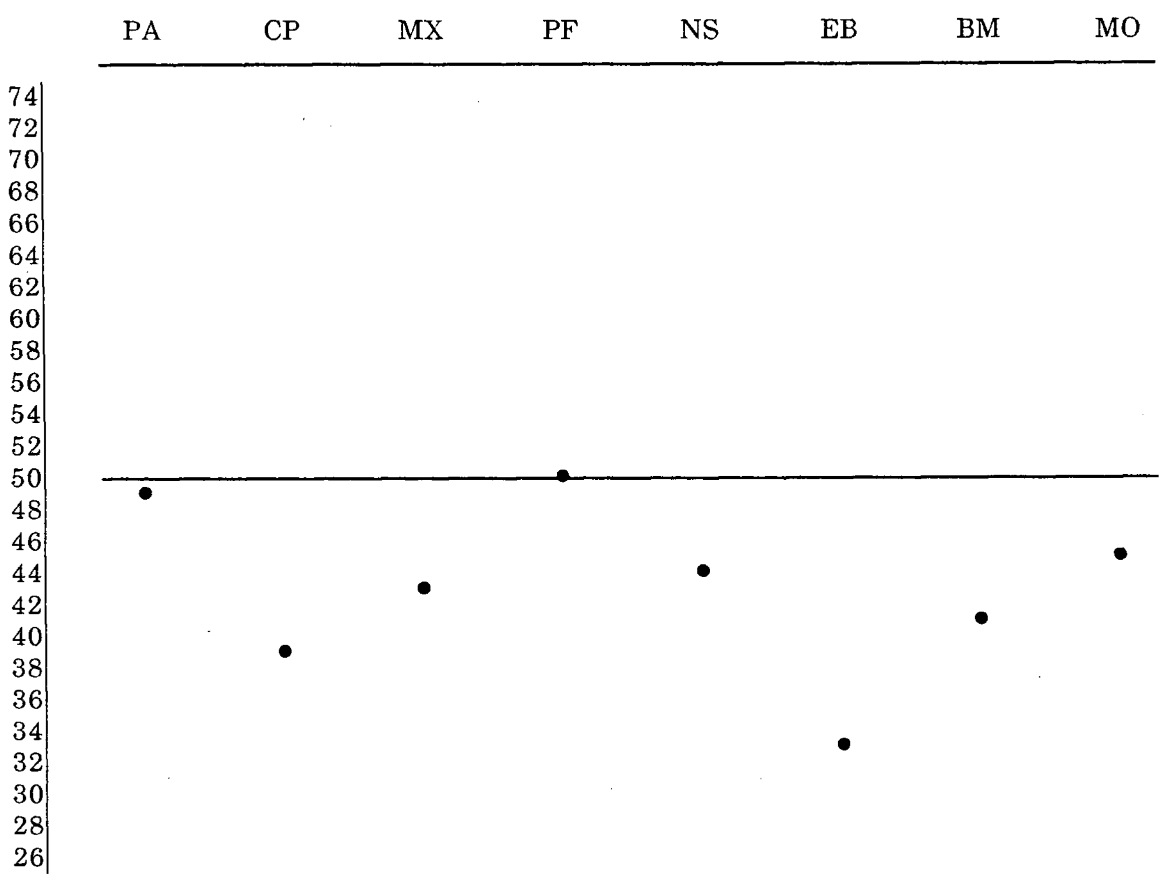

Figure 2

SB:FE Subtest Profile[1]

[1]Mean of 50, standard deviation of 8

Note: PT: Pattern Analysis
CP: Copying
MX: Matrices
PF: Paper Folding & Cutting
NS: Number Series
EB: Equation Building
BM: Bead Memory
MO: Memory for Objects

reason for this is that there is such an emphasis on language acquisition and development that attention to details such as neatness of work presentation may not be accorded the same emphasis as in a class of non-exceptional students. There is also the possibility that, although the study group did not include children with obvious motor disabilities, there may have been children with minor fine motor disabilities. The research literature does indicate that the population of hearing impaired children contains a higher incidence of motor dysfunctions due to the etiologies of the hearing impairment (Vernon, 1968).

For the remaining subtests, the possibility of there being common underlying abilities that account for the depressed scores still remains. The results of the exploratory research questions investigating the correlations between both these intelligence measures and academic measures provide the first indication of what this underlying characteristic may be. Both instruments had significant correlations with the Reading and Number Concepts of the SAT-HI. The WISC-R PIQ correlated significantly with more of the SAT-HI measures than did the SB:FE. However the most striking difference between these two instruments and their correlations with the academic measures was found with the Language measures of the SAT-HI. The largest coefficients were produced between the SB:FE and the Language measures of the SAT-HI. The Partial Composite of the SB:FE and Language produced a correlation of .55. The Quantitative Area correlated .57 with Language, The Short-Term Memory Area yielded the highest correlation of .64 with Language. This particular finding correlating short-term memory skills and language supports work done by Watson *et al.* (1982) who suggested that visual short-term memory is important in English language learning in hearing impaired children.

These findings for the SB:FE are contrasted with the complete absence of significant correlations between this Language measure and any of the WISC-R PIQ subtests. This may be an indication of differences between these two instruments for this hearing impaired sample. The WISC-R PIQ does not correlate with Language whereas the SB:FE does. The question that now arises is how is language involved on the SB:FE?

To answer this question the nature of the task demands of these two instruments should be examined. The Examiner's Handbook of the SB:FE (Delaney and Hopkins, 1987) provides a description of the abilities and influences for each test. While these are considered arbitrary designations and no attempt is made by the authors to claim that all examinees will employ all potential approaches when solving a given task, these listed abilities and influences nevertheless reveal information about the nature of the subtests. The abilities and influences mentioned for the Abstract/Visual subtests involve skills in visual analysis, visual imagery, spatial visualization, planning ability and inductive reasoning. As well, performance may be subject to the influence of flexibility. For the Quantitative subtests, performance is thought to reflect skills in numerical fluency, math concepts and computation, and inductive reasoning. Performance is also thought to be influenced by flexibilty. For the Short-Term Memory subtests, performance is thought to reflect skills in visual memory; visual perception; visual analysis; sequencing, chunking, or clustering strategies; and a verbal labeling memory strategy. These abilities and influences are relevant for a non-exceptional population. Whether they are also relevant to a hearing impaired population is not yet known. However the results from this study's investigation into the pattern of subtest intercorrelations (Hypothesis 2) indicate that the SB:FE subtests are interrelating in a similar manner for both populations. It appears that the hearing impaired children were not substituting discernably different strategies.

Considering the descriptions of the abilities required by the subtests of these two instruments leads to a suggestion that while there are similar abilities involved in the visual perceptual area, the SB:FE also seems to involve more thinking skills involving planning ability, strategy formation, inductive reasoning skills and flexibility. The time factor element may be an important distinction between these two tests. While both tests require visual perceptual skills and problem solving abilities, the time element may suggest different problem solving strategies must be employed for either test. The WISC-R Performance Scale is designed to reward immediate problem solving skills. The SB:FE, considered a power test by its authors, places more emphasis on problem solving strategies employing more reflective thinking skills. The inductive reasoning required by the SB:FE seems to require a sequential ordering of information. The reader is cautioned however, that because of the limited sample size available for the SAT-HI Language measures, rather than being regarded as empirical support, the results have served to initiate speculation regarding the association of language and the SB:FE.

This suspected association found further support with the WISC-R Performance Scale and SB:FE subtest intercorrelations. Drawing on the results of the exploratory research questions involving the subtest intercorrelation of both the instruments provides another clue to the possible nature of these underlying abilities. The subtest intercorrelations analysis found Picture Completion, Block Design, and Picture Arrangement, to yield the highest median correlations of .64,

.63, and .61, respectively, with the Stanford-Binet: Fourth Edition subtests. This indicates that these WISC-R Performance subtests share the most commonality with the Stanford-Binet: Fourth Edition subtests. The correlational analysis with the SAT-HI academic measures indicated that Picture Completion and Picture Arrangement were most highly correlated with academic achievement. It is interesting to note that in a review of factor analysis of the WISC-R for the non-exceptional population, Sattler (1988) stated that Picture Completion and Picture Arrangement had moderate loadings on the Verbal Comprehension factor. This was interpreted to suggest that these two subtests may require verbal mediation to a greater degree than do the other Performance subtests. This provides an insight for a possible explanation for the commonality shared by these two subtests, the Stanford-Binet subtests, and higher language achievement in the hearing impaired group. The use of language as a cognitive tool may be the most common factor shared in these different cognitive tasks. These indications appear to suggest that these selected nonverbal performance type subtests of the SB:FE reflect a model of cognition that links cognitive functioning with language use. That cognition may be facilitated by language is certainly not a novel idea in the study of the nature of the relationship of language and cognition. Vygotsky (1962) has produced a framework for the study of internalized, private speech and its role in directing cognition. Vygotsky sought to trace the development of speech from its initial interpersonal use to its eventual intrapersonal use. Luria (1961) also believed that internalized speech comes to control thought and action. From Luria's research it seems apparent that people who develop a verbal language also develop verbal language mediation strategies to facilitate cognition. The use of language as a cognitive tool in devising memory strategies or in the inductive reasoning process may well be the common underlying factor necessary for successful execution of these subtests. This is not to suggest that hearing impaired individuals do not use language as a cognitive tool. Indeed, the nature of the systems used by the hearing impaired population for symbolic mediation is an active area of research. A discussion of this field is deferred to the directions for further research section. At this point the most important issue is that the standardization population has access to verbal language based cognitive tools that may not be equally available to most deaf children. This leads to the suggestion that the level of language development may have a strong influence in these particular subtests. It could be that cognition and linguistic levels may not be independent in these subtests. It is accepted that most hearing impaired children may be considered delayed in their language development (Quigley and Paul, 1984; Kretschmer and Kretschmer, 1978). Thus in measures where language and cognition are thought to interact, the hearing impaired children would display a lower performance level when compared with their non-exceptional age peers where standard levels of language development may be more readily assumed. This suggestion is reinforced by the method the SB:FE uses to determine entry level into the different subtests. The SB:FE utilizes a multistage adaptive testing pattern. In the first stage, the examiner gives the Vocabulary Test which serves as a routing test. The Vocabulary Test level attained and the chronological age are used to determine the entry level for all other subtests. Within the organization and structure of the SB:FE there is in effect a screening process whereby a child is determined to have a certain level of language before attempting certain subtests.

For the purposes of this study, it is thought that there is enough evidence to suspect that these nonverbal selected subtests of the SB:FE nevertheless require a facility with language as a cognitive tool. For the hearing impaired sample, the tests results from this instrument may therefore be more a reflection of language and/or symbolic mediation abilities than of the underlying cognitive structures. The low scores reported for this hearing impaired group were obtained despite very conservative selection of SB:FE subtests. The battery chosen for this study did not include several of the subtests recommended in the SB:FE Examiner's Handbook for a deaf examinee. Several recommended subtests were excluded due to their overt language or reading requirements. This study's battery more closely resembled the battery recommended for examinees considered to have Limited English Proficiency or who are Non-Language Proficient. This finding therefore, would suggest even greater concern over the suitability of the recommended battery for Deaf Examinees and may also have implications for other populations of exceptional examinees.

C. CONCLUSIONS

The results of this study suggest the following conclusions regarding the appropriateness of the nonverbal selected subtests of the SB:FE for this population of severe to profoundly hearing impaired children whose educational methodology is Total Communication.

1. Results from the selected subtests of the SB:FE and from the WISC-R Performance Scale cannot be used interchangeably. The very significant mean differences between the scores of these two instruments for this population of hearing impaired children preclude their use together as confirmatory estimates of cognitive ability.

2. There are indications that the selected subtests of the SB:FE may not measure cognitive structures independently from language. Therefore, further research using the SB:FE is necessary before it is used as an indicator of general cognitive ability for hearing impaired children whose levels of language development may be delayed.

D. LIMITATIONS OF THE STUDY

The primary limitation of this study may be considered typical of research with special low incidence populations. A balance, or compromise must be made between the number of population variables chosen for control, and the number of subjects available for the research study. The population variables selected for control in this study, as explained in Chapter III, were considered the most important for the nature of this study. The population used in this study is thought to be representative of most populations of hearing impaired children attending a day/residential school whose educational methodology is Total Communication. The results of this study can therefore only be regarded as having implications for similar populations.

Another limitation of the study is the use of modified administration procedures with standardized tests. This issue represents another necessary compromise in research populations with special needs. The effects of the modified procedures were minimized by conferring with the school's Sign Consultant for suitability of the signs used as well as the consistency across both tests. Nevertheless, such test administration modifications would introduce an unknown degree of variation.

Yet another limitation of this study is the very small sample size available to calculate the SAT-HI Language correlations. The seven students who completed this test were the older students in the study. This would certainly limit the generalizability of these results.

E. DIRECTIONS FOR FURTHER RESEARCH

One of the findings of this study indicated the selected subtests of the SB:FE may have predictive validity for this hearing impaired sample. The comparatively high correlations produced between this instrument and the academic achievement measures needs to be verified. If subsequent research confirms this finding, then an understanding of how language is involved in the execution of these subtests may provide information that might be incorporated into educational practices. Certainly, language learning is a central focus in the education of deaf children.

The SB:FE seems to require the use of language-based cognitive strategies. Whether these strategies are based in verbal language (English) or whether these cognitive strategies may be based in a form of Sign Language needs to be investigated. The high correlation between English language and the SB:FE need not imply a direct causal relationship. The work of Cummins (1984) may have implications for this investigation. Cummins proposed a model that suggested the level of language proficiency in the first language is directly related to proficiency in the second language. The applicability of Cummins' framework could be verified using carefully defined subgroups of hearing impaired children. The most important variable in this investigation would be level of first language development. If it is found that a hearing impaired subgroup apparently uses cognitive strategies as effectively as their age appropriate peers in the non-exceptional population, further investigations could probe the exact nature of the cognitive strategies. It is possible that some cognitive strategies may employ a verbal language based system in the Short-Term Memory subtests. Conrad (1979) has conducted studies indicating that some deaf people use speech coding in short term memory tasks.

It is also conceivable that there are other cognitive strategies necessary for the execution of the subtests requiring certain problem solving skills. It has been noted that several of the subtests require inductive reasoning. As this reasoning process seems to require the deliberate, sequential progression of logical statements, it could be that a system of symbolic mediation is employed. Research in this area is growing. King and Quigley (1985) have concluded that the many studies in this area demonstrate a growing belief that a gestural form of language, such as American Sign Language, is probably an efficient thought-mediating system. The suggestion that a reflective, sequential progression of logical thoughts characterize the inductive reasoning processes required by several of the SB:FE subtests leads to yet another interesting speculation. Perhaps the effects of the delay of language development is related to the development of successive cognitive processes as suggested by the model of cognitive functioning proposed by Das, Kirby, and Jarman (1979).

An age appropriate internalized language system may have significant academic and behavioural consequences. It is anticipated that these problem solving skills may be of greater academic consequence in the higher grades and at the post secondary levels. As well, the ability to anticipate possible outcomes and modify or choose alternate behavioural responses may be facilitated by an internalized language-based mediation system.

To the extent that such an internalized mediation system is important in acquiring certain problem solving skills, the the SB:FE may prove to be important in providing an estimate of these abilities.

REFERENCES

Allen, W., Abraham, S., & Stoker, R. (1988). Providers and practices in psychoeducational assessment of the hearing impaired in educational settings. Diagnostique, 14(1), 26-48.

Amoss, H. (1949). Ontario School Ability Examination. Toronto: Ryerson Press.

Anastasi, A. (1982). Psychological testing (5th ed.). New York: Macmillan.

Anderson, R.J. & Sisco, F.H. (1977). Standardization of the WISC-R Performance Scale for deaf children (Series T.No.1). Washington, D.C.: Gallaudet College Office of Demographic Studies.

Arthur, G. (1955). Arthur Adaption of the Leiter International Performance Scale. New York: The Psychological Corporation.

Baker, C. & Cokely, D. (1980). American Sign Language: A teacher's resource text on grammar and culture. Silver Spring, Maryland: T.J. Publishers.

Braden, J.P. (1985). WISC-R deaf norms reconsidered. Journal of School Psychology, 23, 375-382.

Bragman, R. (1982). Review of research on test instructions for deaf children. American Annals of the Deaf, 127, 337-346.

Brill, R.C. (1962). The relationship of Wechsler IQs to academic achievement among deaf students. Journal of Exceptional Children, 28, 315-321.

Brooks, C.R., & Riggs, S.T. (1980). WISC-R, WISC, and reading achievement relationships among hearing-impaired children attending public schools. The Volta Review, 82(2), 96-102.

Burgemeister, B.B., Blum, L.H., & Lorge, I. (1972). Columbia Mental Maturity Scale. New York: Harcourt Brace Jovanovich, Inc.

Cummins, J. (1984). Bilingualism and special education. San Diego: College-Hill Press Inc.

Conrad, R. (1979). The deaf schoolchild. London: Harper & Row.

Das, J.P., Kirby, J.K., & Jarman, R.F. (1979). Simultaneous and successive cognitive processes. New York: Academic Press.

Delaney, E.A. & Hopkins, T.F. (1987). Examiner's handbook: An expanded guide for fourth edition users. Chicago: The Riverside Publishing Co.

Evans, L. (1980). WISC Performance Scale and Coloured Progressive Matrices with deaf children. British Journal of Educational Psychology, 50, 216-222.

Federal Register. (August 23, 1977). Public Law 94-142 (p. 42494).

Furth, H. (1966). Thinking without language: Psychological implications of deafness. New York: Free Press.

Galan, A.R.N., & Nelson, L.R. (1986). LERTAP 3 technical reference manual. North Vancouver, B.C.: Ditactics Services.

Gannon, J.R. (1981). Deaf heritage: A narrative history of deaf America. Silver Spring, MD: National Association of the Deaf.

Gardener, E., Rudman, H., Karlsen, B., & Merwin, J. (1982). Stanford Achievement Test (7th ed.). San Antonio, TX: Psychological Corp.

Gerweck, S., & Ysseldyke, J.F. (1975). Limitations of current psychological practices for the intellectual assessment of the hearing impaired: A response to the Levine study. The Volta Review, 77(4), 243-248.

Glass, G.V., & Hopkins, K.D. (1984). Statistical methods in education and psychology (2nd ed.). New Jersey: Prentice-Hall, Inc.

Goetzinger, M.R., & Houchins, R.R. (1969). The 1947 colored Raven's Progressive Matrices with deaf and hearing subjects. American Annals of the Deaf, 11(4), 95-101.

Goodenough, F.L., & Harris, D.B. (1963). Goodenough-Harris Drawing Test. New York: The Psychological Corporation.

Graham, E.E., & Shapiro, E. (1953). Use of the Performance Scale of the Wechsler Intelligence Scale for Children with the deaf child. Journal of Consulting Psychology, 17(5), 396-398.

Hirshoren, A., Hurley, O.L. & Hunt, J.T. (1977). The WISC-R and the Hiskey-Nebraska test with deaf children. American Annals of the Deaf, 122(4), 392-394.

Hirshoren, A., Kavale, K., Hurley, O.L. & Hunt, J.T. (1977). The reliability of the WISC-R Performance Scale with deaf children. Psychology in the Schools, 14(4), 412-415.

Hirshoren, A., Hurley, O.L., & Kavale, K. (1979). Psychometric characteristics of the WISC-R Performance Scale with deaf children. Journal of Speech and Hearing Disorders, 44(1), 73-79.

Hiskey, M.S. (1966). Hiskey-Nebraska Test of Learning Aptitude. Lincoln, NB: College View Printers.

James, P.R. (1984). A correlational analysis between the Raven's Matrices and WISC-R Performance Scales. The Volta Review, 86(7), 336-341.

Kaufman, A., & Kaufman, N. (1983). Kaufman Assessment Battery for Children. Circle Pines, MN: American Guidance Service.

King, C.M. & Quigley, S.P. (1985). Reading and deafness. San Diego: College-Hill Press.

Kretschmer, R.R. & Kretschmer, L.W. (1978). Language developement and intervention with the hearing impaired. Baltimore: University Park Press.

Kyle, J.G. (1980). Measuring the intelligence of deaf children. Bulletin of the British Psychological Society, 33, 54-57.

Lane, H. (1938). Measurement of the mental and educational ability of the deaf child. Journal of Exceptional Children, 4, 169-173.

Leiter, R.G. (1948). Leiter International Performance Scale. Chicago: Stoteling Co.

Levine, E.S. (1981). The ecology of early deafness: Guides to fashioning environments and psychological assessments. New York: Columbia University Press.

Levine, E.S. (1974). Psychological tests and practices with the deaf: A survey of the state of the art. The Volta Review, 76, 298-319.

Levine, E.S. (1971). Mental assessment of the deaf child. The Volta Review, 73(2), 80-105.

Livingston, S. (1983). Levels of development in the language of deaf children: ASL grammatical processes, Signed English structures, semantic features. Sign Language Studies, 40, 193-286.

Luria, A.R. (1961). The role of speech in the regulation of normal and abnormal behavior. New York: Liveright.

McQuaid, M.F. & Alovisetti, M. (1981). School psychological services for hearing impaired children in the New York and New England area. American Annals of the Deaf, 126(1), 37-42.

Moores, D., Weiss, K., & Goodwin, M. (1976). Recommended policies and procedures: Preschool programs for hearing-impaired children. Minneapolis: University of Minnesota Research, Development and Demonstration Center in Education of Handicapped Children, Research Report No. 104.

Murphy, L.J. (1957). Tests of abilities and attainments. Pupils in schools for the deaf aged six to ten. In A.W.G. Ewing (Ed.), Educational guidance and the deaf child (pp. 213-251). Manchester, England: Manchester University Press.

Myklebust, H.R. (1960). The psychology of deafness. New York: Grune & Stratton.

Myklebust, H.R. (1966). The psychology of deafness, (2nd ed.). New York: Grune & Stratton.

Neuhaus, M. (1967). Modifications in the administration of the WISC-R Performance Subtests for children with profound hearing losses. Exceptional Children, 33, 573-574.

Phelps, L. & Ensor, A. (1986). Concurrent validity of the WISC-R using deaf norms and the Hiskey-Nebraska. Psychology in the Schools, 23(2), 138-141.

Pintner, R., Eisenson, J., & Stanton, M. (1941). The psychology of the physically handicapped. New York: Crofts.

Porter, L.J. & Kirby, E.A. (1986). Effects of two instructional sets on the validity of the Kaufman Assessment Battery for Children - nonverbal scale with a group of severely hearing impaired children. Psychology in the Schools, 23(1), 37-43.

Quigley, S.P. (1969). The influence of fingerspelling on the development of language, communication, and educational achievement in deaf children. Urbana, IL: Institute for Research in Exceptional Children.

Quigley, S.P. & Paul, P.V. (1984). Language and deafness. San Diego, California: College-Hill Press.

Raven, J.C. (1960). Standard Progressive Matrices. London: H.K. Lewis & Company.

Raven, J.C. (1965). The Coloured Progressive Matrices Test. London: H.K. Lewis & Company.

Ray, S. (1979). An adaptation of the Wechsler Intelligence Scale for Children-Revised for the deaf. Natchitoches, LA: Northwestern State University of Louisianna.

Reed, M. (1970). Deaf and partially hearing children. in P. Mittler (Ed.), The psychological assessment of mental and physical handicaps. London: Methuen.

Ritter, D.R. (1976). Intellectual estimates of hearing-impaired children: A comparison of three measures. Psychology in the Schools, 13(4), 397-399.

Rosenstein, J. (1961). Perception, cognition, and language in deaf children. Exceptional Children, 27, 276-284.

Salvia, J. & Ysseldyke, J.E. (1985). Assessment in special and remedial education (3rd ed.). Boston: Houghton Mifflin.

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 (3rd ed.). San Diego: Jerome M. Sattler.

Snijders, J.T., & Snijders-Oomen, N. (1970). Non-verbal Intelligence Test for Deaf and Hearing Subjects (4th ed.). Groningen, Holland: Wolters-Noordhoff.

Spragins, A., & Gibbins, S. (1982, March). Interaction effect: School psychology and deafness in training and practice. Paper presented at the meeting of the National Association of School Psychologists, Toronto, Ontario.

SPSS Inc. (1988). SPSS-X user's guide. Chicago: SPSS Inc.

Steiger, J.H. (1979). MULTICORR: a computer program for fast, accurate, small-sample testing of correlational pattern hypotheses. Educational and Psychological Measurement, 39, 677-680.

Stokoe, W.C. (1970). Sign language diglossia. Studies in Linguistics, 21, 27-41.

Sullivan, P.M. (1982). Administration modifications on the WISC-R Performance Scale with different categories of deaf children. American Annals of the Deaf, 127(6), 780-788.

Sullivan, P.M. & Vernon, M. (1979). Psychological assessment of hearing impaired children. School Psychology Digest, 8(3), 271-290.

Terman, L., & Merrill, M. (1973). Stanford-Binet Intelligence Scale, 1972 norms edition. Boston: Houghton Mifflin.

Thorndike, R.L., Hagen, E.P., & Sattler, J.M. (1986a).Guide for administering and scoring the SB:FE. Chicago: The Riverside Publishing Co.

Thorndike, R.L., Hagen, E.P., & Sattler, J.M. (1986b). Stanford-Binet Intelligence Scale: Fourth Edition. Chicago: The Riverside Publishing Company.

Tweney, R.D., & Hoemann, H.W. (1976). Translation and sign language. In R.W. Brislen (Ed.), Translation: Applications and research. New York: Gardner Press, Inc.

Ulissi, S.M. & Gibbons, S. (1984). Use of the Leiter International Performance Scale and the Wechsler Intelligence Scale for Children - Revised with hearing impaired children. Diagnostique, 9, 142-153.

Vernon, M. (1976). Psychological evaluation of hearing impaired children. In L. Lloyd (Ed.), Communication assessment and intervention, (pp. 195-224). Baltimore: University Park Press.

Vernon, M. (1968). Fifty years of research on the intelligence of deaf and hard-of-hearing children: A review of literature and discussion of implications. Journal of Rehabilitation of the Deaf, 1(4), 1-12.

Vernon, M. (1967). Relationship of language to the thinking process. Archives of General Psychiatry, 16, 325-333.

Vernon, M., & Brown, D.W. (1964). A guide to psychological tests and testing procedures in the evaluation of deaf and hard-of-hearing children. Journal of Speech and Hearing Disorders, 29(4), 414-423.

Vygotsky, L.S. (1962). Thought and language. Cambridge, MA: MIT Press.

Watson, B.U. & Goldgar, D.E. (1985). A note on the use of the Hiskey-Nebraska Test of Learning Aptitude with deaf children. Language, Speech, and Hearing Services in Schools, 16, 53-57.

Watson, B., Goldgar, D.E., Kroese, J., & Lotz, W. (1986). Nonverbal intelligence and academic achievement in the hearing impaired. The Volta Review, 88(3), 151-158.

Watson, B., Sullivan, P.M., Moeller, M.P., & Jensen, J.K. (1982). Nonverbal intelligence and English language ability in deaf children. Journal of Speech and Hearing Disorders, 47(2), 199-204.

Wechsler, D. (1949). Wechsler Intelligence Scale for Children. New York: The Psychological Corporation.

Wechsler, D. (1974). Wechsler Intelligence Scale for Children-Revised. New York: The Psychological Corporation.

Woodward, J. (1973). Some characteristics of Pidgin Sign English. Sign Language Studies, 3, 39-46.

Zimmerman, I.L. & Woo-Sam, J. (1972). Research with the Wechsler Intelligence Scale for Children: 1960-1970. Psychology in the Schools, 9, 232-271.

APPENDIX A

ITEM ANALYSIS CHARACTERISTICS OF SELECTED SUBTESTS OF THE SB:FE

Table A1

Item Analysis Characteristics of SB:FE Subtest: Pattern Analysis

Item	% Excluded	diff	discr.
13	0.0	97.4	.13
14	0.0	0.0	.00
15	0.0	97.4	.31
16	0.0	0.0	.00
17	0.0	97.4	.13
18	0.0	0.0	.00
19	0.0	0.0	.00
20	0.0	94.7	.46
21	0.0	97.4	.33
22	0.0	92.1	.49
23	2.6	97.3	.37
24	2.6	94.6	.41
25	2.6	70.3	.54
26	2.6	67.6	.60
27	5.3	72.2	.51
28	5.3	83.3	.58
29	15.8	68.8	.69
30	15.8	75.0	.81
31	28.9	77.8	.76
32	28.9	48.1	.58
33	39.5	69.6	.73
34	39.5	78.3	.85
35	47.5	95.0	.86
36	47.5	75.0	.75
37	50.0	52.7	.72
38	50.0	68.4	.74
39	57.9	66.7	.77
40	57.9	50.0	.58
41	68.4	66.7	.66
42	68.4	75.0	.65

Table A2

Item Analysis Characteristics of SB:FE Subtest: Copying

Item	% Excluded	diff	discr.
14	0.0	94.8	.09
15	0.0	65.8	.43
16	0.0	92.1	.40
17	0.0	55.3	.57
18	0.0	68.4	.37
19	7.9	71.4	.51
20	7.9	60.0	.50
21	21.1	53.3	.44
22	21.1	66.7	.82
23	26.3	39.3	.59
24	26.3	42.9	.71
25	52.6	38.9	.54
26	52.6	55.6	.72
27	68.4	33.3	.52
28	68.4	41.7	.54

Table A3

Item Analysis Characteristics of SB:FE Subtest: Matrices

Item	% Excluded	diff	discr.
1	0.0	100.0	.00
2	0.0	78.9	.63
3	0.0	78.9	.62
4	0.0	68.4	.51
5	15.8	71.9	.66
6	15.8	93.8	.55
7	18.4	67.8	.78
8	18.4	54.8	.58
9	28.9	48.1	.81
10	28.9	55.6	.72
11	47.4	50.0	.68
12	47.4	70.0	.83
13	63.2	78.6	.84
14	63.2	42.9	.59
15	65.8	76.9	.81
16	65.8	61.5	.67
17	71.1	54.5	.71
18	71.1	63.7	.66
19	73.7	50.0	.66
20	73.7	40.0	.59
21	84.2	66.7	.63
22	84.2	0.0	.00
23	86.8	20.0	.37
24	86.8	20.0	.26
25	89.5	0.0	.00
26	89.5	0.0	.00

Table A4

Item Analysis Characteristics of SB:FE Subtest: Paper Folding & Cutting

Item	% Excluded	diff	discr.
1	2.6	73.0	.42
2	2.6	37.8	.56
3	2.6	21.6	.82
4	2.6	32.4	.56
5	47.4	35.0	.86
6	47.4	30.0	.84
7	78.9	87.5	.87
8	78.9	37.5	.35
9	78.9	2.5	.15
10	78.9	37.5	.76
11	84.2	33.3	.75
12	84.2	50.0	.65
13	94.7	100.0	.75
14	94.7	50.0	.61
15	94.7	50.0	.61
16	94.7	100.0	.75
17	94.7	100.0	.75
18	94.7	50.0	.61

Table A5

Item Analysis Characteristics of SB:FE Subtest: Number Series

Item	% Excluded	diff	discr.
1	0.0	86.8	.60
2	0.0	92.1	.54
3	0.0	71.1	.53
4	0.0	81.6	.63
5	10.5	76.5	.54
6	10.5	76.5	.60
7	15.8	96.9	.72
8	15.8	46.9	.69
9	15.8	53.1	.67
10	15.8	71.9	.78
11	31.6	46.2	.73
12	31.6	42.3	.68
13	55.3	94.1	.81
14	55.3	41.2	.60
15	60.5	66.7	.68
16	60.5	60.0	.67
17	63.5	0.0	.00
18	65.8	7.7	.28
19	86.8	20.0	.28
20	86.8	60.0	.46
21	97.4	0.0	.00
22	97.4	0.0	.00
23	97.4	0.0	.00
24	97.4	0.0	.00
25	97.4	0.0	.00
26	97.4	0.0	.00

Table A6

Item Analysis Characteristics of SB:FE Subtest: Equation Building

Item	% Excluded	diff	discr.
1	18.4	0.0	.00
2	18.4	51.6	.62
3	18.4	6.5	.76
4	18.4	22.6	.83
5	84.2	66.7	.91
6	84.2	83.3	.93
7	86.8	80.0	.87
8	86.8	60.0	.86
9	86.8	0.0	.00
10	86.8	60.0	.81
11	92.1	0.0	.00
12	92.1	0.0	.00

Table A7

Item Analysis Characteristics of SB:FE Subtest: Bead Memory

Item	% Excluded	diff	discr.
1	0.0	100.0	.00
2	0.0	97.4	.48
3	0.0	100.0	.00
4	0.0	100.0	.00
5	0.0	100.0	.00
6	0.0	100.0	.00
7	0.0	97.4	.15
8	0.0	100.0	.00
9	0.0	94.7	.41
10	0.0	97.3	.48
11	0.0	89.5	.42
12	0.0	73.7	.28
13	2.6	86.5	.06
14	2.6	86.5	.24
15	2.6	86.5	.40
16	2.6	86.5	.33
17	2.6	81.1	.22
18	2.6	83.8	.38
19	2.6	78.4	.47
20	2.6	59.5	.55
21	7.9	54.3	.41
22	7.9	59.5	.49
23	23.7	31.0	.69
24	23.7	56.6	.58
25	44.7	71.4	.76
26	44.7	38.1	.52
27	57.9	31.3	.45
28	57.9	56.3	.59
29	68.4	66.7	.74
30	68.4	41.7	.68
31	78.9	87.5	.72
32	78.9	25.0	.52
33	84.2	33.3	.52
34	84.2	16.7	.31
35	92.1	66.7	.47
36	92.1	100.0	.62
37	92.1	33.3	.31
38	92.1	0.0	.00
39	94.7	0.0	.00
40	94.7	0.0	.00

Table A8

Item Analysis Characteristics of SB:FE Subtest: Memory for Objects

Item	% Excluded	diff	discr.
1	0.0	100.0	.00
2	0.0	100.0	.00
3	0.0	100.0	.00
4	0.0	76.3	.59
5	0.0	73.7	.60
6	0.0	68.4	.60
7	7.9	22.9	.56
8	7.9	20.0	.63
9	34.2	24.0	.59
10	34.2	16.0	.57
11	86.8	60.0	.65
12	86.8	40.0	.60
13	92.1	0.0	.00
14	92.1	0.0	.00

APPENDIX B

COMPARISONS OF SB:FE TEST SCORES USING ALTERNATE SCORING

Table B1

Comparisons of SB:FE Scores Using Alternate Scoring*

Student	Age	Sex	WISC-R PS	SB:FE (1)	diff	SB:FE (2)	diff
1	11-1	b	112	94	-18	95	-17
2	11-8	b	88	74	-14	72	-16
3	13-1	g	77	78	1	85	8
4	10-8	b	104	101	-3	102	-2
5	13-6	b	106	75	-31	72	-33
6	14-6	b	112	94	-18	95	-17
7	11-10	b	91	83	-8	82	-9
8	16-7	g	90	78	-12	81	-9
9	9-3	b	133	123	-10	123	-10
10	9-3	b	138	122	-17	123	-16
11	16-8	b	101	79	-22	81	-20
12	16-8	g	46	54	8	56	10
13	9-1	g	96	91	-5	91	-5
14	12-8	b	92	75	-17	78	-14
15	16-7	g	80	68	-12	68	-12
16	15-0	g	106	87	-19	88	-18
17	16-9	g	85	71	-14	70	-15
18	12-8	g	112	115	3	119	7
19	14-7	g	112	111	-1	114	2
20	12-5	b	146	124	-22	122	-24
21	14-4	b	100	69	-31	65	-35
22	11-6	b	81	77	-4	70	-11
23	15-6	g	73	74	1	69	-4
24	8-3	b	129	104	-25	103	-26
25	14-6	b	78	69	-9	69	-9
26	15-6	g	118	87	-31	92	-26
27	14-7	b	105	88	-17	96	-9
28	15-6	b	128	82	-46	83	-45
29	14-0	b	80	64	-16	62	-18
30	8-3	b	98	93	-5	85	-13
31	16-7	g	87	71	-16	74	-13
32	9-6	b	95	67	-28	63	-32
33	12-4	g	120	92	-28	97	-23
34	16-2	g	108	89	-19	90	-18
35	15-9	b	96	72	-24	69	-27
36	9-7	g	87	91	4	88	1
37	10-9	b	95	86	-9	83	-12
38	13-6	b	90	64	-26	64	-26

* Alternate scoring as explained in Chapter 3.