Standardization of Stanford-Binet Intelligence Scale-Fifth Edition (Translation and Adaptation in Bangla for Use in Bangladesh)

Nigar Sultana

Session: 2009-10

Department of Special Education
Institute of Education and Research
University of Dhaka

Standardization of Stanford-Binet Intelligence Scale Fifth Edition

(Translation and Adaptation in Bangla for Use in Bangladesh)

Nigar Sultana

Registration Number: 48 (Re-add)

Session: 2009-10

Department of Special Education

Institute of Education and Research

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Special Education

University of Dhaka

Submission Page

The undersigned certifies that she has read a thesis entitled **"Standardization of Stanford-Binet Intelligence Scale Fifth Edition(Translation and Adaptation in Bangla for Use in Bangladesh)"** submitted by **Nigar Sultana** in partial fulfillment of the requirements for the Degree of Doctor of Philosophy and recommends to the University of Dhaka for its acceptance.

Sharmin Huq

Sharmin Huq

Supervisor

20.05.12

Date

Table of Contents

List of Tables

List of Figures

List of Equations

ACKNOWLEDGEMENTS

Standardizing the Stanford-Binet Intelligence Scale is a surprisingly complex and lengthy process requiring years of effort by many people. Since the beginning of this study in 2005, evidence is abundant that host of people contributed directly or indirectly to its successful completion.

First and foremost the author wishes to extend her heartfelt thanks to Dr. Sultana S. Zaman, Professor Emeritus, Institute of Education and Research (IER), University of Dhaka who is the founder of Special Education Department as well as a pioneer in the field of test development in Bangladesh . Of all the persons, I wish to express my deepest appreciation to Dr. Sharmin Huq, Professor, Department of Special Education, Institute of Education and Research (IER), University of Dhaka who served as the supervisor for this research study. Her invaluable suggestions, spontaneous assistance, helpful guidance and continued encouragement and tolerance have made this research successful.

Besides, I express my sincere thanks to Dr. Shaheen Islam, Professor, Department of Educational and Counseling Psychology and Dr. Shirin Zaman Munir, Department of Psychology, University of Dhaka for their valuable advice. Several individuals specially Dr.Md Ilyas, Professor, Department of Psychology, helped for statistical guidance and Mr.Nasir Uddin Khan, Research Associate (BRAC) , Mr. Mahmood Sarwar Shojol and Ms. Dilruba Khanam deserve special thanks for assisting me in the statistical analysis and computer programming of the data and write up format in APA 6th Edition, who spent long hours with much patience to help me in this regard. The author acknowledges the valuable advice and support of her colleagues from the Institute of Education and Research, University

of Dhaka. Besides, I would also like to acknowledge the valuable support of my Provost and House Tutors at Begum Fazilatunnessa Mujib Hall.

I would like to express my sincere thanks to all the institutions and the children aged 6 to 16 years who participated for the study. The author would also like to thank a number of former students of the Department of Special Education, IER, who worked as test-administrators for testing the children. The author also would like to extend her sincere thanks to BANBEIS for providing the list of schools and Dutch-Bangla Bank Foundation to grant financial support for the research.

Finally, the author wishes to give her heartfelt thanks to the members of her family, specially her parents, husband and her children for their constant encouragement, assistance and forfeit support during the whole period of the present study.

Nigar Sultana

PhD Research Fellow

Department of Special Education

Institute of Education and Research (IER)

University of Dhaka

ABSTRACT

Children differ qualitatively from their peers in respect to their intellectual abilities. These qualitative differences may influence a child's subsequent independence in his/her life as well as family and society. But it is unfortunate for those parents whose expectations and hopes are shattered by the birth of children at high risk and/or children with developmental delays. Psychologists and educators are systematically utilizing scientific methods to measure individual differences among people. Bangladesh has a dreadful need to improve and update the standard of existing assessment condition which is an integral part of instruction, as it determines whether or not the goals of education are being met. The study aims to standardize the Stanford-Binet Intelligence Scale (Fifth Edition) in Bangla for use in urban Bangladesh in order to fill up the gap in the psychometric sector. Hence, the research was designed to complete the criterion for standardizing the psychological ability test. Thus, the present research was conducted in four steps (item analysis, norm development, reliability and validity) as a part of standardization process of an intelligence scale. For the calculation of norm, the study has considered students from six divisional metropolitan cities to represent Bangladesh. After translating the original SB5 into Bangla, item analysis, as a first step of standardization process, was carried out through SB5 tool kit among the 330 students of 11 age levels (6-16 years) to scrutinize the strengths and weaknesses of the test items. In order to retain the original theme, the items were replaced with native content/symbol or object, made the items culture friendly, and often retranslated the questions for better understanding of the students. The overall reliability coefficient (α=0.84) suggests that there is high and increasing correlation among the items. Based on the raw scores obtained from the ten subtests, age norm was calculated separately for the 11 age groups. The norms were developed on 3300 students from the raw scores that were obtained in the record form. Their raw scores were

ranked in 19 scores group and then the Full Scale Intelligence Quotient (FSIQ) was constructed from the ranks. The IQ ranges of SB5-BD for age norm of 6 to 16 years children are 86 to 152. Test-retest reliability was constructed based on the scores obtained twice with the same instrument on the same individual with one week of time interval on 330 students. Test statistics suggests that there was no significant difference between the IQ obtained in the first week and again second administration that was obtained a week later. As a measure of reliability, the correlation coefficient between the first and second administration of the tests were 72%, 76%, and 75% for Non verbal IQ, Verbal IQ and Full Scale IQ respectively. In order to examine the criterion related validity SB5-BD and WISC-R (Bangla Version) were administered on the same participants. The study considered 90 students from three age groups. Findings reveal from the descriptive statistics that there were significant similarities between the IQ scores obtained by the two test administration. To find out the differences in IQ for test validity, the SB5-BD was administered on normal and students with special needs. Result indicates a low mean and standard deviation among students with special needs. The P value suggests that there is statistically significant difference between the IQ obtained by normal and students with special needs. Finally, the study extensively accomplished the four steps and standardization process successfully completed. Through this study, the standardized SB5-BD is regarded as the renovative and contemporary assessment scale in the field of psychometric testing for 6 to 16 years children in urban Bangladesh and hope; it will accelerate all the stagnant issues related to the benefit of human kind, above all for the children with special needs.

ABBREVIATIONS/SYMBOLS

16PF	Sixteen Personality Factor
A1	First Test Administration
A2	Second Test Administration
AERA	American Educational Research Association
AYP	Annual Year Plan
APA	American Psychological Association
ATVR	Analogies Test of Verbal Reasoning
BANBEIS	Bangladesh Bureau of Educational Information and Statistics
BD	Bangladesh
BKT	Binet Kamat Scale of Intelligence
CD	Compaq Disk
CHC	Cattell-Horn-Carroll
CI	Confidence Interval
CRC	Convention on the Rights of the Child
D	Discrimination Index
DAS	Disability Assessment Schedule
DAT	Differential Aptitude Test
DDST	Denver Developmental Screening Test
DEO	District Education Officer
DPE	Directorate of Primary Education
DPI	Disabled People International
DSM-IV-TR	Diagnostic and Statistical Manual, Fourth Edition, Text Revision

DSS	Department of Social Services
EFA	Exploratory Factor Analysis
EFA	Education for All
EPT	Educational and Psychological Testing
FR	Fluid Reasoning
FSIQ	Full Scale IQ
ICDDR, B	International Centre for Diarrhoeal Disease Research, Bd
ID	Intellectual Disability
IE	Inclusive Education
IER	Institute of Education and Research
IEP	Individualized Education Program/ Individual Education Plan
IQ	Intelligence Quotient
IRT	Item Response Theory
ITC	International Test Commission
K-BIT	Kaufman Brief Intelligence Test
K/Kn/KN	Knowledge
M	Mean
MAT	Miller's Analogy Test
MDG	Millennium Development Goal
MMY	Mental Measurements Yearbook
NCLBA	No Child Left Behind Act
NCTB	National Curriculum Textbook Board
NFR	Nonverbal Fluid Reasoning
NK	Nonverbal Knowledge
NCME	National Council on Measurement in Education

NQR	Nonverbal Quantitative Reasoning
NGO	Non Government Organization
NVSP	Nonverbal Visual Spatial Processing
NWM	Nonverbal Working Memory
P	Difficulty Index
PCSAS	Partial Composite Standard Age Score
PEDP-II	Primary Education Development Project- II
P.L.	Public Law
QR	Quantitative Reasoning
RAD	Rapid Assessment of Disability
RIAS	Reynolds Intellectual Assessment Scales
SAS	Scaled Age Score
SB1	Stanford Revision and Extension of the Binet - Simon Scale.
SB2	New Revised Stanford-Binet Tests of Intelligence
SB3	Stanford-Binet Intelligence Scale, Third Revision
SB4	Stanford-Binet Intelligence Scale: Fourth Edition
SB5	Stanford Binet Intelligence Scale Fifth Edition
SB5-BD	Stanford Binet Intelligence Scale Fifth Edition Adapted in Bangladesh
SD	Standard Deviation
S-FRIT	Slosson Full-Range Intelligence Test
SNE	Special Needs Education
UNESCO	United Nations Education Scientific and Cultural Organization
VFR	Verbal Fluid Reasoning
VK	Verbal Knowledge

VQR	Verbal Quantitative Reasoning
VSP	Visual-Spatial Processing
VVSP	Verbal Visual Spatial Processing
VWM	Verbal Working Memory
W.H.O	World Health Organization
WISC-III	Wechsler Intelligence Scale for Children - Third Edition
WJ III COG	Woodcock-Johnson III Tests of Cognitive Abilities
WM	Working Memory

GLOSSARY OF PSYCHOMETRIC TERMS

Ability

Ability describes the degree to which someone can carry out certain types of psychological or mental reasoning operations.

Ability Test

Ability test refers to a test designed to measure intelligence, aptitude or achievement. While aptitudes are defined as inherent abilities for learning and raw talents, they are sometimes measured through a person's achievements because the tests are designed incorrectly. The many types of ability tests include intelligence, verbal, numeric, literacy and abstract reasoning tests.

Aptitude Test

Standardized tests measuring specific intellectual capabilities or other characteristics.

Aptitude is a potential to succeed at something and aptitude tests are designed to measure those mental abilities which affect the likelihood of someone acquiring some particular skill.

Cognition

The conscious process of knowing or being aware of thoughts or perceptions, including understanding and reasoning.

Cognitive Abilities Test

The cognitive Abilities test (CogAT, CAT) is a K-12 assessment designed to measure students' learned reasoning abilities in the three areas most linked to academic success in school: Verbal, Quantitative and Nonverbal. Although its primary goal is to assess students' reasoning abilities, CogAT can also provide predicted achievement scores when administered with the Iowa Tests. The author of the test is David F. Lohman of the University of Iowa.

CogAT is also often used to help educators make student placement decisions, especially when selecting students for Gifted and Talented programs.

Composite Test Score

A score produced by adding together scores of two or more tests which may be differentially weighted before they are added. Composite scores derived from batteries of tests are often produced to provide a general measure of suitability in a selection situation.

Confidence Interval

The score range that can be predicted with a certain level of confidence (usually 95%) for the same child sitting the same test again.

Correlation

A statistical relationship between two variables such that high scores on one factor tend to go with high scores on the other factor (positive correlation) or that high scores on one factor go with low scores on the other factor (negative correlation).

Criterion Referenced Test

A test taker's score is used to predict how they will perform on types of task not directly sampled by the test but which have been shown to be correlated with performance in that test.

Crystallized Intelligence

It is accumulated information absorbed from culture and life experiences. It includes the application of skills and knowledge to solving problems. Crystallized intelligence is ability that is dependent on acquired knowledge.

Domain

A domain refers to the class of ability or achievement. For example, verbal reasoning, numerical reasoning, spatial ability etc.

Descriptive Statistics

Data summarized in numerical form, such as mean, median, mode. This forms the first stage of data analysis. Means, standard deviations and standard errors are presented in the form of a table.

Disability

Disability summarizes a great number of different functional limitations occurring in any population, in any country of the world. People may be disabled by physical, intellectual or sensory impairment, medical conditions or mental illness. Such impairments, conditions or illnesses may be permanent or transitory in nature. (United Nations. (1993). The Standard Rules on the Equalization of Opportunities for Persons with Disabilities.)

Dropout

Children with disability enrolled in different education system but left school after a certain time for different reasons and not continuing to study any more.

Enrolment

Children with disability enrolled in to different educational systems being practiced in different countries.

Factor

A term which usually refers to the independent variable. If two different independent variables are used they are referred to as Factor A and Factor B. Each factor may have different levels.

Factor Analysis

Factor analysis is a statistical technique used to isolate underlying relationship between sets of variables. Factor analysis searches for such joint variations in response to unobserved latent variables. Factor analysis originated in psychometrics, and is used in behavioral sciences, social sciences, and other applied sciences that deal with large quantities of data.

Fluid Intelligence

It is a kind of raw learning ability or the capacity to reason in an abstract way. It includes the speed with which information can be analyzed and also includes attention and memory capacity. Fluid intelligence is natural ability that is not dependent on acquired knowledge.

Flynn Effect

A rise in IQ of the general population of about 3 points per decade, discovered by New Zealander, James Flynn in the early 1980s. When the new test subjects take the older tests, in almost every case their average scores are significantly above 100. To compensate for the IQ increase, test makers select a new sample for the norm reference on their tests about every ten years.

Formal Education

Formal education is defined as "the institutionalized, hierarchically structured, chronologically graded education system starting from primary to post-primary levels of education" (BANBEIS, 1999).

Full-Scale IQ (FSIQ)

FSIQ is the sum of all the tasks on an IQ test – all subtests covering both the Verbal and Non-verbal domains of cognitive ability. It is often referred to as a global summary of the child's current level of intellectual functioning. The FSIQ is usually the most reliable score available from the SB5 or WISC4 because it uses all facets of the test.

Global or general ability: 'g' factor

Spearman noted the positive correlation among the various tasks on IQ tests. He named this the 'g' factor. Global ability is the composite of abilities which enables an individual to learn and recall information, communicate with others, recognize likeness and differences, reason quantitatively and to apply these abilities in solving problems and dealing effectively with the

environment. The *g* factor is represented numerically by the full-scale IQ score on the SB-5 and WISC 4.

General Ability Tests

These tests vary from those designed to give an overall measure of general intellectual ability through assessing broad areas of ability (for example, verbal reasoning or numerical reasoning) to those focusing on specific mental operations (for example, three-dimensional spatial rotations). General ability tests tend to include items dealing with each of the main areas of ability.

Generalization

An inference made from a sample to a population. The researcher attempts to extend the results of his/her study to a much larger group of people.

Inclusive Education

Inclusive education means that… schools should accommodate all children regardless of their physical, intellectual, social, emotional, linguistic or other conditions. This should include disabled and gifted children, street and working children, children from remote or nomadic populations, children from linguistic, ethnic or cultural minorities and children from other disadvantaged or marginalized areas or groups. (The Salamanca Statement and Framework for Action on Special Needs Education, Para 3)

Integrated Education

The pedagogic concept of integration is that, " … it involves the admission of children with special educational needs in 'ordinary' or 'regular' schools and may be described as 'pedagogic integration'. This may be mandatory under legislation, or it may take the form of statements of policy which aim to encourage such integration" (UNESCO 1996). Educational integration refers to measures taken to provide education within the regular education system

with some extra support (i.e. resource room, resource teacher etc.) for children with special educational needs.

Item Response Theory

In psychometrics, Item Response Theory (IRT) is a paradigm for the design, analysis, scoring of tests, questionnaires, and similar instruments measuring abilities, attitudes, or other variables. This theory focuses on the item by modeling the response of an examinee of given ability to each item in the test.

Mean

A measure of central tendency, giving an average of a set of scores (i.e. the sum of all the scores divided by the number of scores in the set).

Median

Measure of central tendency, giving the value of the middle most score (above or below which half of all the scores lies). If there is an even number of scores the median is the average of the two middle scores.

Norm

A statistical concept in psychometrics representing the aggregate responses of a standardized and representative group is established for a test, against which a subject is compared.

Norms

Information usually in the form of a table, which enables raw scores to be converted into standard scores or percentile scores.

Norm Group

The sample of people from whom norms are derived.

Normal Distribution

The normal distribution (a bell-shaped curve) represents a theoretical frequency distribution of measurements. In a normal distribution, scores are concentrated near the mean and

decrease in frequency as the distance from the mean increases. The mean, mode and median are all equal to each other; the proportion of the values falling between any interval along the scale is known from the mathematical properties of the distribution. There will always be, for example, 68% of the values between -1 and +1 standard deviations.

Norm Referenced Test

This defines where a test taker's raw score lies in relation to the scores obtained by the norm group.

Normative Sample

Norm-referenced tests are expected to assess skills that are representative of the children in the relevant population. Because the test cannot be standardized on the entire relevant population - which is often very large - the test developers select a much smaller normative sample that is expected to represent that population.

Percentile

The value on the raw score scale below which a given percentage of the sample's scores lie. For example, if the 75th percentile rank is 40, then 75% of the sample will have scored less than 40.

Population

Psychometrics involves making inferences about people who come from some population on the basis of information known about the behavior of a representative sample from that population.

Potential

A capacity to perform or acquire the skills to perform some class of actions.

Psychological test

A psychological test is essentially an objective and standardized measure of a sample.

Psychometrics

Psychometrics is the quantitative and technical aspect of intellectual measurement.

Rasch models

Rasch models are used for analyzing data from assessments to measure variables such as abilities, attitudes and personality traits. Rasch models are particularly used in psychometrics, the field concerned with the theory and technique of psychological and educational measurement.

Raw Score

The raw score of an individual is the initial scores given based on his or her correct responses to test items. Typically, the raw scores of different individuals on most standardized tests are not meaningfully compared with each other.

Reliability

The consistency with which a measuring instrument (such as a psychometric test) performs its' function, gauged, for example, by comparing test scores from the same subjects at different times.

Repetition

Children with disability or special needs enrolled in different education system but could not get promotion to the higher classes and retained in the same class for 2/3 years.

Sample

A subgroup selected from a larger group of potential subjects (population).

Scaled Score

Most standardized test scores are statistically transformed into other kinds of scores, collectively known as derived scores or scales. Such derived (transformed) scores of different individuals may be compared to determine their relative strengths and limitations.

Percentiles, standard scores (z), and stanines are among the more frequently encountered transformed scores in standardized test manuals.

Source Language

The source language in a test adaptation process refers to the language from which the test is being translated and adapted. When translating an instrument from language A to language B, language A is the source language.

Special Education

Every child is unique, and every child needs help in developing and adjusting to life. Some children need more help than other. And some need special help over longer or shorter periods of their lives, for example, during their school years. This special, extra help is often referred to as 'Special Education'.

Special Needs Education

"The term 'special needs education' has come into use as a replacement for the term 'special education'. The older term was mainly understood to refer to the education of children with disabilities that takes place in special schools or institutions distinct from, and outside of, the institutions of the regular school and university system. In many countries today a large proportion of disabled children are in fact educated in institutions within the regular system.

Standard Deviation

A measure of dispersion within a set of data, calculated from the square root of the variance, to give a value in the same range as raw scores. The standard deviation is the spread of scores around the mean of the sample.

Standardization

Standardization is a research process that includes careful selection of test items, administration of the items to a representative sample drawn from a defined population,

statistical analysis of results, establishment of age-based norms, and development of instructions and response scoring procedures.

Standard Scores

The standard scores represent the degree to which a child's score deviates from the mean. Deviation from the mean is expressed in terms of a standard deviation (SD); it is a measure of the distance between the group mean and an individual score. There are two common types of standard scores: the Z score and the T score.

Standardized Tests

Standardized tests are those that have been administered to a large group of individuals who are similar to the group for whom the test has been designed.

Stanford-Binet (SB5)

Psychometric assessment tool which must be administered by an educational psychologist. A measure of a child's global intellectual functioning (or *g*) is derived from the composite of subtests. Recommended for ages 2 to 85.

Stratified Sampling

The most accurate way of developing norm group. Test developer as well as users take into account all demographic variables which can accurately describe the population of interest and then selects individual at random, but proportional to the demographic portrait of the test population. Common demographics to stratify: age, gender, socioeconomic status, geographic region.

Target Language

Target Language is the language of the new instrument. When translating an instrument from language A to language B, language B is the target language.

Test Administrator

A test administrator is the person who administers a psychometric test.

Testing

Take measures to check the quality, performance, or reliability of something (e.g. intelligence),

Testlets

Groups of three to six items combine to form testlets in SB5 test kit.

Test–retest Reliability

Test–retest reliability scores on the test administration to the same persons are correlated with those obtained on separate occasions.

Test Standardization

Test standardization is the psychometric process followed in the test construction of standardized tests used in testing.

Validation

The process of building up evidence about what can and cannot be inferred from test scores.

Validity

From the Latin validus, (strong), the degree to which a measuring instrument measures what it is supposed to measure. Information on the validity of a test tells the user what inferences can be drawn about the person who has produced the score on a test and what is being measured by a test.

Weighted mean

The weighted mean is similar to an arithmetic mean (the most common type of average), where instead of each of the data points contributing equally to the final average, and some

data points contribute more than others. The weighted arithmetic mean is used, if one wants to combine average values from samples of the same population with different sample sizes. If all the weights are equal, then the weighted mean is the same as the arithmetic mean.

Wechsler Intelligence Scale for Children 4 (WISC-4)

Psychometric assessment tool which must be administered by an educational psychologist. A measure of a child's global intellectual functioning (or *g*) is derived from the composite of subtests. Recommended for ages 6 to 17.

Z-score

A standard score scale with a mean of zero and a standard deviation of one. A score expressed in units of standard deviations from the mean. Also known as a standard score.

CHAPTER ONE

Introduction

The supreme adaptive resource of human being is his intelligence – his superior intellectual ability for wisdom, interpretation and prediction. By the blessings of this resource, as a species, he dominates on many facets of his environment and establishes his superiority over other associates of the living kingdom. Besides, this supremacy may be flattened and ruin the individual's spontaneous autonomy if an individual is born with or acquire developmental delays. The pursuit of an efficient and accurate way to identify and compare this ability in individual is an ongoing trends and its consequence in the field of education and development is apparent and undeniable. Thus, scholars of the earlier period explored intelligence to categorize the individual differences and their abilities but there were variations among the experts in defining intelligence in a single concept.

Hence, along with the above perspective, the present study would be considered as the Hallmark reformation in the field of educational development in Bangladesh. Consequently, the study is an efficient and renovative effort to establish a yardstick for the benefit of human kind, above all children with special needs. This chapter briefly presents five issues: firstly, it states the understanding the concepts of intelligence and individual differences ; describes the psychometric tests, secondly, it explains the necessity of assessing intelligence, thirdly, the chapter highlights the psychometric and contemporary device for intelligence test, fourthly, it portrays the present testing and disability scenario: international and Bangladesh perspective, finally , this chapter depicts the rationale and objectives of the study.

Understanding the Concepts– Intelligence and Individual Differences

Intelligence is versatile and often changed notion with referred to as Intelligence Quotient (IQ), cognitive functioning, intellectual ability, and aptitude, thinking skills, general ability and intellectual development (Logsdon, 2011). These multifaceted terminologies are being used throughout the study to comprehend the unique criteria of intelligence. Everyone assumes that he or she knows intelligent performance when it is observed, but when it is tried to define, the ambiguity of the trait becomes apparent (Daniels, Devlin & Roeder, 1997). With various common consents of researchers', numerous definitions of intelligence have been proposed before the twentieth century. Besides, various approaches to human intelligence also have been adopted of which few have been explained to validate the present research study.

The unitary concept of general ability or intelligence emerged from the definitions of Binet and Spearman. In their studies, they created a statistical technique called factor analysis to explore their approach. From the studies, they were able to report that about half of the variance in tests of mental ability was due to the general factor (Kaplan & Sacuzzo, 2001). This general or global intelligence is commonly referred to by the single italicized letter, *g* (Spearman, 1927). An alternative conception of intelligence is that cognitive capacities within individuals are a manifestation of a general component, or general intelligence factor, as well as cognitive capacity specific to a given domain such as reading, mathematics and writing (Miller, 1991). Even though at present intelligence is viewed in a multidimensional concepts as emotional, multiple, social, artificial intelligence etc. In this study, the author intends to utilize the general intelligence as a global perspective to justify an individual's intellectual capabilities that influence his /her overall developmental condition particularly academic and social performance.

The concept of individual differences was gaining popularity around the world at the same time as Binet's work, spurred by the movement towards universal compulsory education in many countries. At the time, many psychologists were addressing the problem of how to identify children who would have success in education (Thorndike, 1990). Thinking on the same aspect on intelligence, the pioneer of intelligence testing, Binet (1905) reflected the opinion that *"In intelligence there is a fundamental faculty, the alteration or the lack of which, is of the utmost importance for practical life. This faculty is judgment, otherwise called good sense, practical sense, initiative, the faculty of adapting one's self to circumstances."* A heightened focus on defining and assessing intelligence began in the 1800's as part of attempts to classify between various levels of mental retardation and mental illness using psychological tests (Anastasi & Urbina, 1997). Viewed broadly, the scientific and professional organization, the American Psychological Association (APA, 1996) defmes intelligence with the concept that *"Individuals differ from one another in their ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought."* These definitions seemed to have an orientation to academic learning and performance along with emphasis on abilities that are valued by one's culture. As cultural differences play a vital role in forming an individual's life style, it is essential to assess how different cultures make sense of the world in terms of the meanings that represent the mind and within which the concept of intelligence is defined (Bouchard & Segal, 1985).

Therefore, at present the most acceptable definition of this concept is *"Intelligence is not a single, unitary ability, but rather a composite of several functions. The term denotes that combination of abilities required for survival and advancement within a particular culture"* (Anastasi, 1992; 1997). Thus, more recent definitions have been moving toward

practical definitions with a view as to how the person functions in the real world as well as in traditional academic settings (Wagner, 2000). Aspects of the definition that seem to have wide appeal include learning speed, adaptability and ability to perform in the society successfully.

Hence, research in intelligence is active as well as robust, and this study investigates the power of intelligence related to educational, social learning and performance of both normal and children with special needs. Further, a great number of researches still have been conducted through various ways, using many theoretical viewpoints and establishing a variety of results to define and measure intelligence throughout the year. Despite the variety of terms of intelligence, the most influential approach to understanding intelligence is based on psychometric testing. In fact, the technical term for the science behind psychological testing is psychometrics (Neisser, Boodoo, Bouchard, Boykin, Brody & Ceci, 1996).

Psychometric Tests

Psychometrics is the field of study concerned with the theory and technique of psychological measurement, which includes the measurement of knowledge, abilities, attitudes and personality traits. The field is primarily concerned with the study of differences between individuals. It involves two major research tasks, namely: (i) the construction of instruments and procedures for measurement; and (ii) the development and refinement of theoretical approaches to measurement. (Kline, 1999).

The first psychometric instruments were designed to measure the concept of intelligence. The best known historical approach involves the Stanford –Binet intelligence scale, developed originally by the French Psychologist Alfred Binet. Contrary to a fairly widespread misconception, there is no compelling evidence that it is possible to measure

innate intelligence through such instruments, in the sense of an innate learning capacity unaffected by experience, nor was this the original intention when they were developed.

Similarly, psychological testing is a field characterized by the use of samples of behavior in order to assess psychological construct(s), such as cognitive and emotional functioning, about a given individual. The burning issue at present in the field of psychology is the assessment (referred to as test, evaluation, measurement, scale, battery etc.) of an individual's behavioral characteristics (e.g. ability of intelligence, emotional functioning, interests or attitudes, aptitude, normal, abnormal personality and achievement) through psychological tests. Psychological assessment is also referred to as psychological testing, or performing a psychological battery on a person. This is also a process of testing that uses a combination of techniques to help arrive at some hypotheses about a person and their behavior, intelligence, personality and capabilities (Framingham, 2011). Assessment can range from the formal--standardized to the informal--teacher made assessments. Standardized tests are usually considered as formal tests. These are developed by testing organizations and administered in clinics and class room settings and scored in a consistent manner. In this aspect, the test scores are interpreted with regards to a norm or criterion, or occasionally both. The norm is established independently, or by statistical analysis of a large number of participants (Mellenbergh, 2008). There are several categories of psychological test, such as achievement test, aptitude tests, intelligence tests, neuropsychological tests, occupational tests, personality tests etc (Charles, 1996).

Table 1

Several Categories of Psychological Tests (At a Glance)

Test name	Setting /Used in	What Measure	Example
Achievement test	Educational	Achieved knowledge	General Certificate of Secondary Education (GCSE) Test of English as a Foreign Language (TOEFL)
Aptitude test	Employment	Aptitude	Scholastic Aptitude Test (SAT)
Intelligence test	Clinic / School	Potential/ Intelligence	WISC-R, SB5
Neuropsychological	Clinic	Deficits in cognitive functioning	Cambridge Neuropsychological Test Automated Battery (CANTAB)
Occupational	School / Office	Interest in career	Occupational Interest Profile
Personality	Forensic	Personality	Minnesota Multiphasic Personality Inventory (MMPI)

Source: Charles (1996).

These psychological tests are often discussed in terms of the dimensions as they measure. They refer to these as dimensions because they are broader than a single attribute or trait level. Often these types of tests measure various personal attributes or traits. (Hersen, 2003). Professionals refer to these tests in various ways. Sometimes they refer to them as tests of maximal performance, behavior observation tests, or self-report tests. Sometimes professionals refer to tests as being standardized or non-standardized, objective or projective. Other times they refer to tests based on what the tests measure. (Rasch, 1980:1960). Even though, from above among the various psychological tests, the study focuses only on a standardized norm-referenced intelligence test for assessing the intellectual ability of an individual. The educational need and advanced educational programs for identifying and

classifying children with limited intellectual abilities and gifted learners has been an important force in the development of psychological tests. These tests also play an especially important role in special education. They can be useful for identifying an expected level of academic performance and also in helping school professionals design Individual Education Plan (IEP) for students with special needs (Sattler, 2001). Thus, the testing movement is the consequence of a need to determine the intellectual, sensory, and behavioral (personality) characteristics in individuals and hence, intelligence as a significant factor could only be established until a person's ability is assessed.

The Necessity of Assessing Intelligence

Assessing intelligence is a complex process but has become an established practice in psychological testing because of its potential effects on individuals' lives. Measures of a child's intellectual abilities are considered one part of what is referred to as the 'Fours Pillars of Assessment'. Along with behavioral observations , interview and informal assessment, intelligence testing provides an assessor with information into a child's overall level of functioning , as well as specific abilities (Sattler, 1992). However, intelligence tests provide information about a child's abilities in two main ways that the above stated other methods do not. Firstly, it provides a standardized or norm referenced framework. Secondly, aptitude test has been found to be correlated with performance in both school and work environments (Sattler, 1992, Anastasi & Urbina, 1997).

Children differ qualitatively from their peers in respect to their intellectual abilities. Besides, these qualitative differences may influence a child's subsequent independence in his/her life as well as family and community. But it is unfortunate for those parents whose expectations and hopes are shattered by the birth of children who are at risk or children with

developmental delays. It is no secret that the number of children with special needs has dramatically increased in the past decade worldwide (Reschly, Tilly & Grimes, 1999). Therefore, comparisons between individuals, as well as intra-individual performances can be made for the purpose of placement or identifying special education needs using these tests. According to Diagnostic and Statistical Manual of Mental Disorders Fourth Edition (Text Revision) and American Psychiatric Association (APA), the aim of assessment is to gain insight into an individual that will aide in the decision making process with regard to screening, problem solving, diagnosis, therapy, rehabilitation, progress evaluation and to gauge the necessity for a complete battery (DSM-IV-TR & APA, 2000). Measuring intelligence is based on the fact that children become more capable mentally as they advance in age. The upper limit is reached in adolescence. Intelligence tests show that intellectual growth is rapid in infancy, moderate in childhood, and slows down in youth (Cahan &Cohen, 1989).

Thus a prerequisite criterion for the placement of such children either in mainstream or special school is to quantify their intellectual level that necessitates the measurement of intelligence through intelligence scale in accordance with their age, and sex. (Neisser, 1998). This comprehensive assessment will assist a professional to justify a child's strength and weakness to overcome his delays. Accordingly, the goal of this research was not to categorize children with a single score but to pinpoint a child's intellectual level along with other multidimensional factors such as age, sex, culture. Most significantly, Binet had the similar notion to identify children in the schools who required special educational needs. His intention was not to use IQ scores as a general device for ranking all children according to intellectual ability (Binet & Simon, 1905). Binet's scale had a profound impact on educational

development throughout the world. However, in spite of its constraints, the educators and psychologists utilized the scale worldwide with its actual value.

Based on the above pragmatic demands it can be traced that assessing intelligence among other individual traits has created an outstanding platform that depicts a person's general level of intellectual capability, which is significant for the life of a human being. Moreover, the success of educational system in advanced countries has been owing to the development and utilization of standardized psychological testing of abilities of students. In this aspect, psychologists and educators are systematically updating and standardizing various psychometric and contemporary tests for the last century to measure individual differences among people.

The Psychometric and Contemporary Device for Assessing Intelligence

Ever since Alfred Binet's great success in devising test to distinguish intellectually challenged children (terminologies used earlier were idiot, moron, imbecile, mentally retarded, mentally handicap, and intellectually disabled, intellectual impairment) from those with behavioral problems, psychometric instruments have played an important part in European and American life. Standardized tests are commonly used for historic, regulatory and practical reasons. A variety of historical trends, actual strengths, educational policies and commonly offered arguments justify the use of standardized tests. Tests are used for many purposes, such as selection, diagnosis and evaluation. Many of the most widely used tests are not intended to measure intelligence itself but closely related to construct scholastic aptitude, school achievement and specific abilities etc. Such tests are especially important for selection, decision and placement purposes (Flanagan, Genshaft & Harrison, 1997). Besides, standardized tests have been historically promoted as "objective" in the sense that the

examiner's biases would not influence the results (Domino, 2000). Moreover, psychologists, clinicians are routinely and traditionally trained in administering standardized tests due to the historic belief that standardized assessment is better because they are more formal and objective than other kinds of assessment, which are often named as "informal," implying "less objective." (Anastasi & Urbina, 1997). Therefore, selecting the most appropriate test for a given child or situation can be a challenging task.

A review of the last 10 years of Mental Measurements Yearbooks (MMY) indicates an increase in the number of intelligence tests that can be used for young children. A few well known individually administered intelligence tests are as follows: Stanford-Binet Intelligence Scales, Fifth Edition (SB5) (Roid,2003),Wechsler Intelligence Scale for Children - Fourth Edition (WISC-IV) (Wechsler, 2004), Slosson Full-Range Intelligence Test (S-FRIT) (Algozzine, Eaves, Mann & Vance, 1993), Kaufman Brief Intelligence Test (K-BIT) (Kaufman & Kaufman, 1993) and Woodcock-Johnson III Tests of Cognitive Abilities (WJ III COG) (Woodcock, McGrew & Mather, 2001), Reynolds Intellectual Assessment Scales (RIAS) (Reynolds, 2003). These tests are being used in evaluating intelligence and /or cognitive abilities in schools as well as assessment centre for identification purposes. In addition to this, the tests are developed for norm on large sample sizes and justify the age appropriate intellectual ability (Chang, 2008).

Researchers have different opinions on using these tests for assessment purposes. Along with varied opinions on the use of tests, the experts' have come to a common consents and supports that the Stanford-Binet Intelligence Scale, Fifth Edition is a sole contemporary device with a rich tradition since its inception in 1905 till date. Through various editions, this assessment scale is being used throughout the world. Other strengths of SB5 include its appealing materials and cognitively appropriate tasks. Besides, psychometric properties of the

test at the school age, and its comprehensive subtests are considered as other strengths to find out children's intellectual development in both verbal and nonverbal domains (Ford & Dahinten, 2005). Bracken and Nagle (2007) also suggested the use of the SB5 to assess the cognitive abilities of children as young as school age due to its superior psychometric and qualitative characteristics. Based on its popularity, usability and standard for intelligence measurement, SB5 is acknowledged and considered as the paramount instrument to serve the purpose of the present research. It is to be mentioned that the American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], (1999) have highly recommended the use of SB5 as the Standards for Educational and Psychological Testing. Though several psychological tests have received prominence, many current innovations were derived only from the Binet-Simon scale. With regard to the current standard for educational and psychological testing, the SB5 has earned a leading position in the field of intellectual assessment. This scale is an individually administered assessment of intelligence and cognitive abilities. The Stanford-Binet Intelligence Scales, Fifth Edition (SB5), a direct descendent of Terman's adaptation of the Binet test developed more than 100 years ago, is used in the educational setting. The SB5 is comprised of five composite factors representing two domains as nonverbal and verbal each having five testlets with a total of ten subtests (Roid, 2003) (reviewed and discussed in the chapter two and three) .

Present Testing and Disability Scenario: International and Bangladesh Perspective

Formal and systematic measurement of intelligence, begun with the French psychologists Binet and Simon at the beginning of the 20th century, heralded the modern era of psychological testing. In subsequent years, tests to measure aptitude, personality and educational achievement were developed. The need to assess various abilities of a large

number of army recruits at the beginning of World War I in 1917 gave a significant boost to psychological testing (Gregory, 2007). In the 21st century, psychological testing is a big trade in developed countries especially in America. There are thousands of commercially available, standardized psychological tests as well as thousands of unpublished tests. Approximately 20 million Americans per year were taking psychological tests (Goldman & Saunders, 1995). Today, psychological testing is a part of the American culture. Psychological tests are in use everywhere. The previous tests are regularly used in the school system as tools in making placement decisions. Current research provides information that supports the relationship between achievement and intelligence tests. One of the most significant and controversial uses of psychological testing in the 21st century has been a result of the 'No Child Left Behind Act' of 2001 (NCLB Act). The NCLB Act contains the strategies for improving the performance of schools—strategies that were intended to change the culture of America's schools by defining a school's success in terms of the achievement of its students (U.S. Department of Education, 2004). While tests have always played a critical role in the assessment of student achievement, the NCLB Act requires that students be tested more often and relies on test scores to make more important decisions than in the past. On the contrary, education reform in the United States since the late 1980s has been largely driven by the setting of academic standards for what students should learn and be able to do. These standards can then be used to guide all other system components. The standards-based reform movement describes for clear, measurable standards for all school students. Expectations are raised for all students' performance. Along with norm-referenced rankings, the performance of all students is expected to be raised. Curriculum, assessments, and professional development are aligned to the standards.

Standards-based school reform has become a predominant issue facing public schools. (Popham, 1999). Besides, the largest Flynn effects appear instead on highly g-loaded tests such as Raven's Progressive Matrices. This test is very popular in Europe; Raven's test plays a central role in recent analyses of the worldwide rise in test scores. (Flynn, 2007). Hence, the Flynn effect is coming to an end, at least in Western Europe. Recent studies in Scandinavia show intelligence test scores plateauing and arithmetic scores dropping. Far from being surprised, Flynn has been expecting as much. Since the social condition varies from country to country, it is significant to underpin the context of the diverse world (Flynn, 2007, & Collingwood, 2008).

In relation to the worldwide present scenario of psychological and other testing, Bangladesh is still left behind in the testing pathways. Until recently, most commonly cited disability prevalence rate has been the World Health Organization (WHO), which estimates that approximately 10% of the world's population suffers from disabilities. In Bangladesh context that estimation would interpret as approximately 15 million people with disabilities based on 15th March, 2011 census. Action Aid Bangladesh based on 5 locations of 4 districts cited that approximately 12 million people (14% of the total population) require some form of immediate service due to disability related issues. (Action Aid Bangladesh, 1996). However, lack of quality data about those with disabilities makes addressing their needs difficult. Besides, according to ICDDR, B and core donor AusAID, *"Unless international development programmes are inclusive of and accessible to persons with disabilities, achieving the UN Millennium Development Goal (MDG) is not possible"*. In assistance with University of Melbourne in 2009, ICDDR, B developed a Rapid Assessment of Disability (RAD) toolkit for use by governments, NGOs and other organizations. This toolkit is easy-to-use, comprehensive way to measure disability prevalence, quality of life, social participation,

access to and effectiveness of related development programs. The toolkit contains a four-part questionnaire in collaboration with Australian and Bangladeshi disability organizations and service providers. (Keeffe, Baker, Booth, Goujon, Edmonds, Huq & Quaiyum, 2011). On the other hand, the WHO has designed a set of Disability Assessment Schedules (known as the WHO-DAS) which have a long series of activity and participation based questions.

Moreover, since the formal or mainstream schools run by the government, do not have overall disability programmes or activities at all, very few NGOs are being set to provide the programmes of identification, assessment, placement and decision making for leveling the degrees and type of the disabilities (Choudhuri, Alam, Hasan & Rashida , 2005). Thus with the above discussion till to date assessment plays a central element in the overall quality of teaching and learning in education. It also serves for the purposes of occupational prognosis, for clinical diagnosis, as well as psychological research and theorizing (Devlin, Feinberg, Resnick, & Roeder, 1997; Herrnstein & Murray, 1994). At the end of 19th century a few psychologists and educators have taken the initiative to standardize and develop non-standardized need based assessment scales which at present is outdated with time.

Therefore, no disability prevalence data, the absence of reliable and consistent data on the magnitude and educational status of children with disabilities makes it difficult for educators, policy-makers and programmers to understand the nature of the problem and identify possible solutions.

Rationale of the Study

Appropriate stimulation in childhood occupies one of the most important platforms that influence normal development. Likewise, children use different modes in making sense of their experience and the world around them. They also acquire set of standard norms, knowledge, skills and attitude which the society demands for their existence. In this context,

education (also called learning, teaching or schooling) in the universal sense is any act or experience that has a formative effect on the intelligence, character or physical ability of an individual. In its practical sense, education is the process by which society deliberately transmits its construct ability, knowledge, skills and values from one generation to another.

Globally, the enactment of legal issues related to compulsory and quality education would ensure a positive and desirable change in all aspects of an individual's development. Based on the philosophy of Public Law 107-110 (2001), No Child Left Behind (NCLB) is a comprehensive plan in USA to reform schools, change school culture, empower parents and improve education for all children as well as improve instruction in high-poverty schools. Further the law ensures that poor and minority children also have the same opportunity as other children to meet the challenging academic standards. This law has brought sweeping changes to education across the world. Moreover, the recent implementation of the No Child Left Behind Act (NCLBA, 2002); the government of the United States mandated that all school-age children be tested for educational progress. In order to execute the mandate of NCLBA, along with the assessment provision, the need for translation and adaptation of test would eventually lead to assess student from multicultural and multilingual context (Allalouf, 2003 & Chang, 1999; Mathews, 2003).

Similarly, with the growing interest in cross-cultural research and evaluation, the interest in testing is not limited only in education but also in other fields. Such as psychological, vocational, career planning, selection and international comparative studies. The result of this interest is a boon for psychometrically equivalent, multi-lingual versions of assessment instruments. With the increasing demand for the use of psychological tests in various cultures and countries, the need for translation and adaptation of the test is of main concern. It is also apparent that the test adaptation is appropriate and significant.

In order to re-affirming the vision of Education For All (EFA), it is stated in the World Declaration made at Jomtien (1990) as: *"All children, young people and adults have the human right to benefit from an education that will meet their basic learning needs in the best and fullest sense of the term"* . With a view to ensure quality education as a human right , assessment should be considered as an important prerequisite to determine a student's ability. It will enable the teachers to gear up and tap each individual's talents and potentialities, so that they can benefit from education and improve their lives and transform to their societies. In accordance with the international commitments and legal acts, Bangladesh government has taken a positive initiative through the National Education Policy 2010 by Ministry of Education. This policy has highlighted the improvement of education system by including students with special needs in mainstream schools. It is unfortunate that in order to maintain the standard of the education system, the policy has not given any emphasis on screening and assessment of students' intellectual ability. It should be mentioned that, according to United Nations Development Programme (UNDP) report in 2011, the ranking status of Bangladesh for literacy is 163 and literacy rate is 55.9 %.

In advanced countries the decision for placement of children in regular classroom or special classes is prioritized through a standardized comprehensive individual assessment of the children's needs. The use of such psychometric tests also facilitate teachers in educational planning by providing approach to determine possible teaching learning strategies, which is regarded as a major initiative in order to ensure the goals for achieving education for all.

Similar to many other low income countries, at present in Bangladesh, there have been no attempts to conduct regular national disability prevalence survey by the national statistical agency, Bangladesh Bureau of Statistics (BBS). The evolution of educational systems for children with special needs started from the introduction of special education in low income

country like Bangladesh a long time ago. Over the time, the concept of disability as a social issue rather than a medical issue has become more understood and therefore the concepts of education systems also have been changed and developed towards as an integrated system and more recently an inclusive system, in accordance with local socio-economic and cultural conditions (Choudhuri et al., 2005).

The study Educating Children in Difficult Circumstances states that 8% of children with disabilities in Bangladesh are currently enrolled in various educational institutions (ESTEEM, 2002). Of these, 55% had physical disabilities, 13% were visually impaired, 12% were hearing and speech impaired and 10% had intellectual disabilities. About 68% of enrolled children with disabilities were in government and private primary schools and 15% were in pre-primary educational settings. About 48% were seeking formal education, 23% were in integrated schools, 15% in special education and 5% in inclusive education. Among enrolled children with mild and moderate disabilities, 79% are enrolled in formal educational settings. Of those with severe and profound disabilities, 83% were enrolled in special education. Nearly, 74% of those who are currently not enrolled in any form of education expressed a keen interest in receiving education (ESTEEM, 2002).

These educational systems are being practiced for children with special needs with few numbers. Likewise, the government's Department of Social Services (DSS) is operating 5 special schools for children with visual impairment, 7 for children with hearing impairment, 1 for children with intellectual disability. The DSS is also operating a total of 64 integrated schools for blind children in 64 districts. NGOs are operating many special and inclusive education centers but there is no reliable data available on the number of schools that they operating (Choudhuri, et al., 2005). Although school enrolment (80%) is increasing at a fast rate, but the enrolment of children with disabilities is extremely low. Children with

disabilities are often marginalized in mainstream schools as a result of negative attitudes towards them. A lack of child-centered approaches in education and the physical inaccessibility of schools are other reasons for low enrolment. In addition, some children with special needs are being enrolled into the mainstream education system by default. Some of them transferred from integrated and or special education systems (primarily visual impaired students) while a few make their way to the mainstream education system directly due to self-initiative and interest. Moreover, there are more than a million primary school-age children with assorted disabilities and disadvantages, but without access to basic education. The major shortcomings are due to the lack of educational reformation, improper implementation of the existing education policy and ignorance of parents. Besides, other barriers of failure in schools and low standard of achievement are due to the lack of proper assessment; counseling and guidance are not offered to students and parents before and during the tenure of their education. Similarly, the high rate of dropout after being enrolled is due to improper use of teaching learning strategies as well as other educational provisions. Even, examination or evaluation system is not suitable for these students. Lack of support systems like; IEP (Individual Education Plan) or provision of extra sessions to cope with the mainstream curriculum is remarkable (Choudhuri et al., 2005). Besides, lack of proper assessments of a student's intellectual capability also plays a significant role in classroom performance as well as to hold on to the retention of students to avail school completion certificates. As a result, the necessity to standardize an appropriate and up to date assessment scale has become essential to mitigate the problem of disability prevalence and the present status of quality education for students with special needs and other marginalized population.

Furthermore, the national curriculum is too 'heavy' for children with different types of disabilities especially children with learning difficulties and intellectual disability. Rather,

there is no uniformed curriculum in the special education system among NGOs, to accommodate different types of these children. Different organizations use different curricula developed by them. To mitigate the above issues as well as the Bangladesh Government's priority to initiate the access of all types of children in mainstream, it eventually establishes a major change in education reformation. From that point of view, the children with disability are brought under inclusive setting without any need based modification in the general curriculum. To address this demand and to ensure quality education curriculum adaptation is of prime importance.

It can be mentioned that children with developmental delay also have sub average intelligence as measured by intelligence tests. With this significant limitation the child shows inability to adapt and carry on everyday life activities such as self-care, socializing, communicating, etc. In general, the greater the severity or limitations of a disability, the greater the emphasis has been on developing functional and life skills rather than setting more academic goals. Students with significant developmental delay represent a special population for whom expectations are generally low. However, like all children, these students have many unique strengths and specific talents in different cognitive factors. It is important to perceive and realize the situation of the individual beyond his/her disability and appreciate each child as an individual.

In consequence with the visions of Education for All, UNESCO in 2010 led a global movement in aiming to meet the learning needs of all children, youth and adults by 2015. Based on this, many significant changes have taken place in the policy and system of education in Bangladesh. Over the years from 70s, a handful of professionals in Bangladesh have taken the front role in developing and standardizing psychometric tests for assessing children's ability. In addition, the National Curriculum Textbook Board (NCTB) currently has

also taken initiative for comprehensive modernized and simplified curriculum modification activities in the area of primary education. But it is a daunting challenge for educators to ensure the students with disability especially children with intellectual disability have access to this general curriculum. So there is a need for scanning approach whether the modification and simplification of the NCTB curriculum project for these students can be used both in inclusive and special schools. Consequently, the factors of SB5 can take a leading role in creating the curriculum more functional and effective for students with all types and degrees of intellectual ability.

Keeping pace and continuance with the above stated goals, the existing measures of assessment techniques need to be updated. In this perspective, the Stanford-Binet Intelligence Scale Fifth Edition (2003) which is used worldwide as a measure for identifying students' academic problem was translated, adapted and standardized for Bangladesh culture. The aim of standardization illustrates that the scale would be a suitable asset to mitigate the sufferings of a large population who are at present being deprived of their basic right to education. In brief, the study was to standardize the latest fifth edition of Stanford Binet intelligence scale to serve the following purposes:

- To diagnose students with developmental disabilities and exceptionalities.
- To guide educators and teachers for psycho educational evaluation for special educational placements.
- To assist psychologists for clinical and neuro- psychological assessment.
- To use for further research on intellectual abilities.

- To incorporate the five factors of SB5 with relevant assistance in the modification of general national curriculum.

The Objectives of the Study

The study aimed to standardize the Stanford-Binet Intelligence Scale (Fifth Edition, 2003) in Bangla for use in Bangladesh. However, the specific objectives, as a part of standardization process are stated as follows:

- To translate and adapt the ten subtests of Stanford-Binet Intelligence Scale for children aged 6 to 16 years.
- To determine the reliability and validity of the adapted versions.
- To develop the norm for Bangladeshi children aged 6 to 16 years.

Following the description and importance of assessing the intellectual ability of an individual in this chapter, chapter two will discuss the literature review compiling the historical studies on intelligence and its assessment along with international and national perspectives on standardization of SB5. Besides, chapter three will describe the methods and methodology involved in standardizing the test. Whereas, chapter four will analyze the results found for the study in Bangladesh. Moreover, chapter five will explain the rationale and justification of the research. Finally, the conclusion and implication and further recommendations for the study will be discussed in chapter six followed by the limitations of the study in the field.

CHAPTER TWO

Literature Review

The chapter focuses on the review of psychometric tests, historical studies on intelligence, its assessment and historical perspectives on intelligence test development, history of the Stanford-Binet and its various editions. This chapter also covers the overview of international and national perspectives on standardization of Stanford-Binet Intelligence Scale, description of standardization process and cross cultural assessment. Moreover, this literature review is an approach to enter into the related field of knowledge and offers an opportunity to enhance the understanding for the accomplishment of a quality study.

Prior to the contributions of many theoretical and practicing psychologists in the early nineteen hundreds, the concept of intelligence as it is understood worldwide today was unknown. Thus, the change in focus began unfolding. From its initial pre-scientific and philosophical roots, the study of intelligence changed drastically (Meloff, 1987).

Review of Psychometric Tests

By the end of the 19th century, people attending scientific or industrial expositions were taking various tests that assessed their sensory and motor skills, the result of which were compared against norms (Anastasi & Urbina, 1997). One active area in the scientific research is the tests of psychological characteristics most commonly, intelligence themselves. Intelligence and the ability to assess it, is considered as an important concept in relation to academic settings. Although many claim that intelligence is defined by what intelligence tests measure, many other theorists and researchers argue that this definition is too circular and narrow. Moreover, scores on intelligence tests are designed to reflect the definitions of

intelligence rather than serve as an exact and unqualified representation of intellectual ability (Gardner, Kornhaber & Wake, 1996). Nevertheless, IQ tests are useful tools for various purposes. Moreover, psychometrics is applied widely in educational assessment to measure abilities in domains such as reading, writing, and mathematics. The main approaches in applying tests in these domains have been Classical Test Theory and the more modern Item Response Theory (IRT) and Rasch measurement models (Kline, 1999). Such approaches provide powerful information regarding the nature of developmental growth within various domains.

Besides, college entrance exams, classroom tests, structured interviews, assessment centers, and driving tests are also psychological tests. On the other ways, many popular psychological testing reference books also classify tests by subject. For example, the *Seventeenth Mental Measurements Yearbook* (Geisinger, Spies, Carlson, & Plake, 2007) classifies thousands of tests into 19 major subject categories like as Achievement, Behavior assessment, Developmental, Education, English, Fine arts, Foreign languages, Intelligence, Mathematics, Miscellaneous (for example, courtship and marriage, driving and safety education, etiquette), Multiaptitude batteries, Personality, Neuropsychological, Reading, Science, Sensor motor (Thompson, 2003). Although some are more typical, all meet the definition of a psychological test. Together, they convey the very different purposes of psychological tests. In the following figure, a continuum of some of the most and least commonly recognized types of psychological tests are shown (Chun, Cobb, & French, 1975).

Table 2

A Continuum of Psychological Tests

More Typical	←	→	Less Typical
	Vocational tests		Road portion of driving test
Personality tests	Interest inventories	Self –scored magazine tests	Structured employment interviews
Intelligence tests	Achievement tests	Classroom quizzes and examination	Assessment centers
	Ability tests		

Source: Chun et al., 1975.

Historical Studies on Intelligence and Its Assessment

During the era of psychometrics, intelligence was thought to be a single, inherit entity. The human mind was believed by some to be a "blank slate" that could be educated and trained to learn anything if taught in the appropriate manner (Sternberg, 2000). However, contrary to this notion, an increasing number of researchers and psychologists now believe that the opposite is true; that is, individuals are born with and possess different levels of ability. The development and use of intelligence tests have been one way that researchers and psychologists have attempted to support their argument. While intelligence is one of the most talked about subjects within psychology, there is no standard definition of what exactly constitutes 'intelligence.' Some researchers have suggested that intelligence is a single, general ability; while other believe that intelligence encompasses a range of aptitudes, skills and talents (Horn & Noll, 1994). The following are some of the major theories of intelligence that have emerged during the last 100 years.

Charles Spearman - General Intelligence.

British psychologist Spearman (1904) described a concept and referred to a general intelligence, or the *g factor*. After using a technique known as factor analysis to examine a number of mental aptitude tests, Spearman explained that scores on these tests were remarkably similar. People who performed well on one cognitive test tended to perform well on other tests, while those who scored badly on one test tended to score badly on others. He concluded that intelligence is general cognitive ability that could be measured and numerically expressed.

Louis L. Thurstone - Primary Mental Abilities.

Psychologist Thurstone (1938) offered a differing theory of intelligence. Instead of viewing intelligence as a single, general ability, Thurstone's theory focused on seven different "primary mental abilities." The abilities that he described were: verbal comprehension, reasoning, perceptual speed, numerical ability, word fluency, associative memory, spatial visualization.

Howard Gardner - Multiple Intelligences.

One of the more recent ideas to emerge is Howard Gardner's theory of multiple intelligences. Instead of focusing on the analysis of test scores, Gardner proposed that numerical expressions of human intelligence are not a full and accurate depiction of people's abilities. His theory describes eight distinct intelligences that are based on skills and abilities that are valued within different cultures. The eight intelligences Gardner described are: Visual-spatial Intelligence, Verbal-linguistic Intelligence, Bodily-kinesthetic Intelligence,

Logical-mathematical Intelligence, Interpersonal Intelligence, Musical Intelligence, Intra personal Intelligence, Naturalistic Intelligence (Gardner, 1983).

Robert Sternberg - Triarchic Theory of Intelligence.

Psychologist Robert Sternberg defined intelligence as "mental activity directed toward purposive adaptation to, selection and shaping of, real-world environments relevant to one's life." While he agreed with Gardner that intelligence is much broader than a single, general ability, he instead suggested some of Gardner's intelligences are better viewed as individual talents. Sternberg proposed what he refers to as 'successful intelligence,' which is comprised of three different factors (Sternberg, 1985). The factors are as follows:

- Analytical intelligence: This component refers to problem-solving abilities.
- Creative intelligence: This aspect of intelligence involves the ability to deal with new situations using past experiences and current skills.
- Practical intelligence: This element refers to the ability to adapt to a changing environment.

Based on above theoretical reviews on intelligence, it states that while there has been considerable debate over the exact nature of intelligence, no definitive conceptualization has emerged. Today, psychologists often account for the many different theoretical viewpoints when discussing intelligence and acknowledge that this debate is ongoing (Horn, 1985).

On the other hand, human being has been fascinated by the noticeable differences in mental capacity that has existed among individuals in society. Ideas relating to intelligence remained a philosophical issue until the late nineteenth century when psychologists began the systematic investigation of intelligence (Thompson, 1984). In 1996, Williams reviewed the

definition of intelligence in his studies that most experts would accept the constructs of goal directed behaviors' that are adaptable across environments. He included in his studies the opinions of experts to define intelligence in two themes that are common to both definitions. The first common theme was focused on the individual learning from experience and the second on the individual's ability to adapt to the environment. In several and similar studies, Chen, 2007; Hale & Jansen, 1994; Myerson, 2003, viewed the processing speed and working memory capacity as the currently predominant integrative constructs for explaining g. Much of the difficulty in developing an adequate intelligence assessment tool is the lack of a consensus definition of what the concept actually represents. Before selecting the task of assessing cognitive abilities, those abilities must be operationally defined. François (1995) stated that in order to make use of what intelligence tests explain us; we must first understand what intelligence is. Through the years, the nature of the types of abilities believed to represent intelligence has taken numerous routes. Even the term intelligence itself has recently taken a back seat to a broader viewpoint involving various cognitive abilities.

Spearman, in 1904, put forth the concept of a 'g' factor, or an overall general intelligence, based on the positive correlations between cognitive tests (Duncan, Seitz, Kolodny, Bor, Herzog, Ahmed, Newell, & Emslie, 2000). He used a factor analysis of many cognitive measures in order to suggest that the main underlying component of these measures was an overall intelligence, or 'g' (Spearman, 1904; Duncan et al., 2000).

In 2002, a study by Ken Richardson on "What IQ Tests Test" describes about how human intelligence should be and whether IQ tests actually measure it and if they don't, what they actually do measure. The study suggests that IQ scores can be described in terms of sociocognitive-affective factors that differentially prepare individuals for the cognitive, affective and performance demands of the test. The paper shows that how such factors can

explain the correlational evidence usually thought to validate IQ tests, including associations with educational attainments, occupational performance and elementary cognitive tasks, as well as the inter-correlation among tests themselves.

Studies on Intelligence Test Development

The study of intelligence and its measurement traces its roots to physicians, educators and psychologists who were deeply involved with population at the extremes of intellectual continuum. Esquirol (1938) and Seguin (1907) were committed to the study of intellectually disabled individuals, and Galton (1884) was fascinated by the mental abilities of geniuses. The separate contributions of these pioneers have been profoundly felt in the field of intelligence testing. It was the innovative research investigations of Binet (1903) who focused on the mental abilities of typical or average children at each age, that have had the longest, lasting and most direct effect on individual intelligence testing as we know it today (Anastasi, 1992).

Esquirol made several important contributions, most notably by distinguishing "between the idiots, whose intelligence does not develop beyond a very low level and the demented person" (Peterson, 1925). This distinction between intellectually disabled and emotional disturbance reflected a vital breakthrough for assessment and indicated the primitive state of the art in the early nineteenth century. Esquirol also described a hierarchy of retardation (or feeble mindedness, as it was known in earlier times) with 'idiots' occupying the bottom rung, followed by "imbeciles" and peaking with "morons" (Peterson, 1925). He was well ahead of his time in concluding that the use of language was the most dependable criterion for inferring a retarded individual's intelligence level. Esquirol (1938) was also credited with developing a precursor of the mental age concept by pointing out that an idiot is

incapable of acquiring the knowledge common to other persons of his own age (Anastasi, 1976). Seguin was heavily influenced in his work with mentally retarded individuals by Itard, of Wild Boy of Aveyron fame. Like Esquirol, Seguin (1907) tried to establish criteria for distinguishing between different levels of retardation, although he focused on sensory discrimination and motor control. Optimism regarding treatment of retarded individuals characterized Seguin's approach and he instituted a comprehensive programme of sense training and muscle training techniques much of which live on in present day institutions for the mentally retarded (Anastasi, 1992: 1976).

Francis Galton (1884) transformed his enthusiasm for gifted men and genius and the study of the genetics of intelligence into the development of what was apparently the first comprehensive individual intelligence test. Galton believed that intelligence must be intimately related to sensory abilities because environmental knowledge comes to us via the senses, he developed a series of tests such as weight discrimination, reaction time, visual discrimination, steadiness of hand, keenness of sight and strength of squeeze. His empirical justification for this test battery came from comparisons between gifted and retarded individuals that, not surprisingly, showed obvious superiority in favour of the gifted (Peterson, 1925). Galton's influence spread far beyond his laboratory as "Galton type tests" were developed throughout Europe and the United States. Cattell (1890) coined the term "mental tests"; Galton's influence was clearly evident in Cattel1's 40-60 minute individual examination, as after-images, colour vision, sensitivity to pain and the like (Peterson, 1925). Cattell elaborated on and improved his mentor's methodology by emphasizing the vital notion that administration procedures must be standardized to obtain results that were strictly comparable from person to person and from time to time (Huq, 1992).

Later a challenge was issued to the Galton view of sensory and motor intelligence from Alfred Binet of France. In collaboration with Simon and Henri (1895), Binet conducted numerous investigations of complex mental tasks rejecting the Galton notion that performance on simple, elementary sensory discrimination and motor co-ordination tasks equates to intelligent behavior. According to Cattell (1976) and Horn & Noll (1997), Stella Sharp (1899) directly compared sensory-discrimination tests with tests of complex mental functions and concluded that the simplest mental processes yielded comparatively unimportant information, whereas the tests of Binet and Henri showed much value in assessing "individual psychical differences". Even though initial reaction to the two studies was predominantly antitesting causing a lack of enthusiasm for the Galton-Cattell as well as the Binet-Henri approach in the United States, the methodology of Binet eventually triumphed first throughout Europe and finally in America (Peterson, 1925).

Interestingly, a research by Jensen (1979) and his students Vernon (1981) has revived the early work of Galton to some extent. Although they confirmed that simple reaction time measures contribute little to variation in intellectual function, these researchers have found substantial relationship between intelligence and complex reaction time over repeated trials of the same task. Thus adaptations of Galton's work might yet be found to impact on objective intellectual assessment in future (Huq, 1992).

History of the Stanford-Binet and Its Various Editions

The most revolutionary contribution of all the theorists of their time was that of Alfred Binet and his young associate Theodore Simon. In 1905, they developed a useful tool to assess general intelligence, which is widely cited as the first major break- through in intelligence testing (Roid, 2003).

Early Work of Binet.

As a member of a French governmental commission working on mental retardation, Binet developed a practical test, sensitive to different levels of cognitive development, which could be given during a clinical interview. Alfred Binet's early work began with intelligence testing, when Binet collaborated with Victor Henri to outline a project for the development of a series of mental tasks to measure individual differences (Binet & Henri, 1895).The tasks were designed to differentiate a number of complex mental faculties, including memory, imagery, imagination, attention, comprehension, aesthetic sentiment, moral sentiment, muscular strength, motor ability and hand-eye coordination.

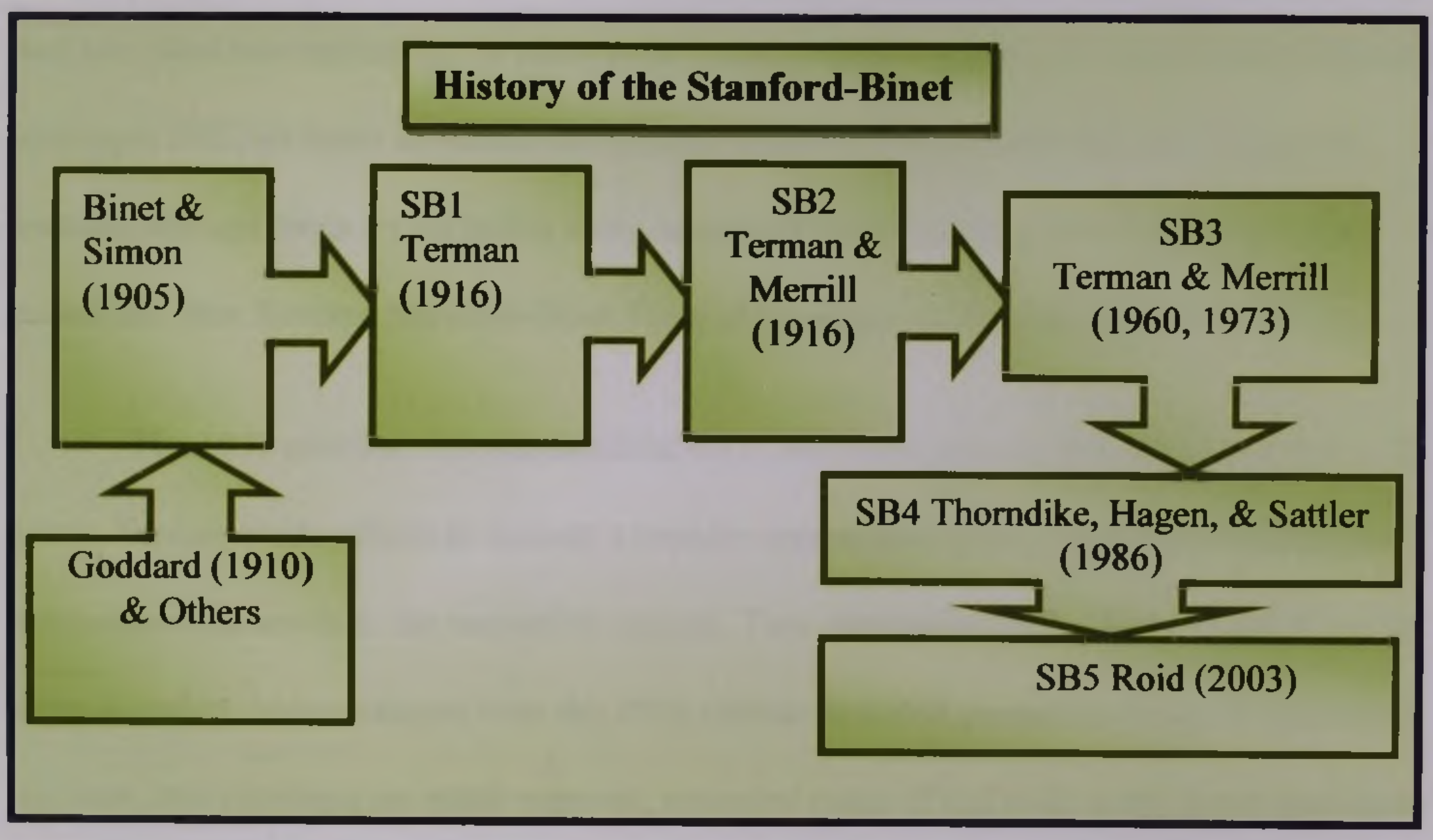
The 1905 Binet - Simon Scale in France.

Binet initiated the leading role in devising a useful and reliable diagnostic system for identifying children with mental retardation. Binet's project culminated in the publication of the first practical intelligence test with physician Theodore Simon (Binet & Simon, 1905). Binet sought to make the 1905 scale efficient and practical: "*We have aimed to make all our tests simple, rapid, convenient, precise, heterogeneous, holding the subject in continued contact with the experimenter, and bearing principally upon the faculty of judgment*" (Binet & Simon, 1916).The scale consisted of 30 items, which were scored on a pass-fail basis. The items presented various word problems, paper-cutting tasks, repeating sentences and digits, and comparing blocks to put them in order by weight (Wolf, 1973).

The 1905 scale included several important innovations that would be used in subsequent measures of intelligence. Items were ranked in order of difficulty and accompanied by careful instructions for administration. Binet and Simon also utilized the concept of age-graded norms (Wolf, 1973).The use of age-graded items allowed the scale to estimate mental age by the pattern of correct answers. The 1905 Binet - Simon Scale was

revised in 1908 (Binet & Simon, 1908) and again in 1911. By the completion of the 1911 edition, Binet had extended the scales through adulthood and balanced them with five items at each age level. The scales included procedures for assessing language, auditory and visual processing, learning, memory, judgment and problem solving (Roid, 2003).

Figure 1. History of the Stanford-Binet.



Source: Becker, 2003.

Terman's 1916 Stanford Revision in America.

Realizing the importance of the theoretical and practical value of Binet's work, Terman (1911) of Stanford University began to adapt the test to the American culture. Within a few years, the improved scale was published as the *Stanford Revision and Extension of the Binet-Simon Scale*. However, Terman's 1916 revision retained Binet's concept of intelligence as a complex mixture of abilities and is the only revision that has stood for

publication to the present day. The standardization that Terman accomplished was quite rigorous for the early 1900s and increased the scale's technical quality (Roid, 2003).

Revisions of the Terman Scales in 1937, 1960 and 1972.

Within 20 years of its release in 1916, the Stanford revision emerged as the most widely used test of intellectual ability in America. The scale had several language translations and was used internationally. In subsequent years, Terman continue to experiment with easier and more difficult items to extend the measurement scale downward and upward and to increase the age range by including more standardization samples. The new edition was called the New Revised Stanford-Binet Tests of Intelligence (Terman & Merrill, 1937).

The 1937 revision was standardized on 3,200 examinees aged 1 year 6 months to 18 years. Terman made efforts to include a broader representation of geographic regions and socioeconomic levels in the normative sample. Two alternative forms, Form L and Form M were included. Improvements over the 1916 edition included greater coverage of nonverbal abilities, less emphasis on recall memory, extended range of the scale at the lower and upper ends, and more objectified scoring methods. (Terman & Merrill, 1937). As happens with any widely used test of ability or achievement, obsolete items were considered for further revision by Terman and Merrill based on the accumulated information and data collected since 1937. Thus, the Stanford-Binet Intelligence Scale, Third Revision, 1960 was published. Several new features were included in the third revision as the use of deviation IQ, (standardized normative mean of 100 and SD of 16), combination of Form L-M while keeping the most discriminating 142 items from the 1937 revision.

After Maud Merrill retired, Robert L. Thorndike of Columbia University was asked to lead a project to collect new norms for the third edition. Thus, the same edition was reprinted

with the new normative tables-an update of Form L-M (Terman & Merrill, 1973). Because the Cognitive Abilities Test (CogAT; Thorndike & Hagen, 1994) was being standardized at the same time as the 1972 reforming of the Stanford-Binet, Thorndike selected subjects and some siblings of subjects tested on the CogAT to compose the new norm sample. The stratification variables used on the sample (e.g., age, geographic region, ethnicity, and community size) was similar to those used today, as were the levels of ability on the verbal portion of the CogAT. The items in the test remained essentially the same as on the 1960 revision, with two minor exceptions.

The 1986 Edition by Thorndike, Hagen and Sattler.

In 1986 Thorndike and his associates accomplished the test with a new appearance and structure. The Stanford Binet Intelligence Scale Fourth Edition (SB4) was based on a four-factor, hierarchical model with general ability (g*)* on a bell curve score (Thorndike, Hagen, & Sattler, 1986). The four cognitive factors were Verbal Reasoning, Abstract/ Visual Reasoning, Quantitative Reasoning and Short-Term Memory. The most significant change from previous editions, however, was the use of point scales for all subtests rather than the developmental age levels used in previous forms. Vocabulary was still retained as a routing test, allowing the test to be tailored to the examinee's verbal ability. Also, many classic Stanford-Binet tasks were retained, including absurdities, vocabulary, matrices, quantitative reasoning and memory for sentences—tasks also included in the SB5. Composite and profile scores for each subtest would permit a comprehensive examination of strengths and weaknesses among abilities within general intelligence (Roid, 2003).

Stanford Binet Intelligence Scale Fifth Edition (SB5) by Gale H. Roid.

Development of the SB5 is heavily based on the new Cattell-Horn-Carroll (CHC) theory of intellectual abilities. In continuation with the past editions as the SB4, five key factors of CHC theory were selected for the development of SB5. In 1995, Gale H. Roid, the author of the SB5 had undertaken the initiative for a new revision and developed it as the fifth edition in 2003. Considering a normative sample of 4,800 subjects, whose ages ranged from 2 to 85 years. The Fifth Edition includes extensive high-end items designed to measure the highest level of gifted performance. It also includes improved low-end items for better measurement and low-functioning of young children with intellectual disability. Furthermore, the inclusion of age-graded norms in SB5 serves as a unique criterion provided for the estimation of mental age (Roid, 2003a; Thorndike, Hagen, & Sattler, 1986).

Composition of the SB5.

The SB5 design crosses the five factors with the two domains resulting in ten (5x2) subtests. Based on the literature such as manuals of SB5 by Roid (2003), the factors, domains and subtests are reviewed below.

Factors.

Factors are the important dimensions of cognitive ability that are measured by the items and subtests of SB5. The factors measured in the SB5 are: Fluid Reasoning (FR), Knowledge (KN), Quantitative Reasoning (QR), Visual-Spatial Processing (VSP) and Working Memory (WM). These factors, the central components of SB5, are discussed below.

Fluid Reasoning (FR)

Fluid Reasoning, as defined by Roid, (2003b) is "the ability to solve verbal and nonverbal problems using inductive or deductive reasoning." The inductive reasoning component requires the individual to derive the general whole from its specific parts. Likewise, the deductive reasoning component requires that the individual draw a conclusion, implication, or specific example from a general piece of information about the topic.

Knowledge (KN)

According to Roid (2003b), knowledge "is a person's accumulated fund of general information acquire at home, school, or work." This construct is often referred to as crystallized intelligence, as it involves learned material that has been stored in long term memory. It also requires perception of detail, attention, concentration, geography, science, and inference skills.

Quantitative Reasoning (QR)

Roid (2003b) defines Quantitative Reasoning, as "an individual's facility with numbers and numerical problem solving, whether with word problems or with pictured relationships" (p. 136). The items included on the SB5 Quantitative Reasoning target problem solving abilities as opposed to rote mathematical knowledge. As the subtests progress, items become more complex.

Visual-Spatial Processing (VSP)

Visual-Spatial Processing, as defined as the "measures an individual's ability to see patterns, relationships, spatial orientations, or the gestalt whole among diverse pieces of a visual display" (p. 137). The items of this factor assess the individual's ability to move pieces

and shapes to form a proper whole. All levels within this area address visual construction abilities (Roid, 2003b).

Working Memory (WM)

In 2003, Roid defines Working Memory, as "a class of memory processes in which diverse information stored in short-term memory is inspected, sorted, or transformed" (p. 137). The individual must filter out the irrelevant information and maintain focus on the pertinent. Furthermore, the information must be manipulated, which places both memory, organizational, and visual-spatial demands on the individual.

Domains.

A domain represents the degree to which a class of item requires the use of language skills, particularly in generating a response to an item. The SB5 contains two domain composites: Nonverbal and Verbal domains. The assessors should consider that the terms "nonverbal "and "verbal" are relative and comparative terms in the SB5. At present, the two domains are discussed accordingly.

Nonverbal Domain

This domain requires less language ability or little or no vocal response or speech and thus has lower language demands. The nonverbal tasks involve a small degree of examiner-spoken directions.

Verbal Domain

This domain requires some degree of expressive language, often as simple as a word or phrase or a degree of reading for the average and high functioning students.

Consideration of the nonverbal versus verbal difference, verbal domain has become increasingly important as society has become more culturally and linguistically diverse.

Subtests.

According to Roid (2003), a subtest is simultaneously an element of exactly one factor and one domain. The ten subtests used in SB5 are described along with figure 2 below along with its contrasting features of the five factors and two domains.

Nonverbal Fluid Reasoning (NFR)

The Fluid Reasoning subtests within the nonverbal domain are Object-Series and Matrices. Initially, the individual is required to match objects. These objects are then placed into a series, either repetitive or not, that the individual must complete. The last phase is similar to the classic matrix-reasoning measures that are common among intelligence testing. (Roid, 2003b).

Nonverbal Knowledge (NK)

The Knowledge subtests within the nonverbal domain include procedural knowledge and picture absurdities. At the lowest end of the spectrum, the subject is required to communicate basic human needs using gesture. As the task demands increase, the subject is presented with impossible pictures in which he is required to point out what is odd or impossible about the scene. The Nonverbal Knowledge tasks tax an individual's basic level of common knowledge about natural phenomena (Roid, 2003b).

Nonverbal Quantitative Reasoning (NQR)

The Quantitative Reasoning subtests within the nonverbal domain have been carried over from the SB4. However, the focus of the subtests from the SB5 is on the reasoning behind the mathematical concepts, as opposed to the rote solving of mathematical items. In order to succeed on the higher level tasks, the subject must use problem solving strategies, persistence, and cognitive flexibility (Roid, 2003b).

Nonverbal Visual Spatial Processing (NVSP)

The Visual-Spatial Processing subtests within the nonverbal domain incorporate the form board activity from the SB4. However, tasks have been added in order to expand the evaluation of Nonverbal Visual-Spatial Processing activities. Initially, shapes are matched and then inserted into forms. As the individual progresses, accurate duplication of patterns using the provided shapes is targeted (Roid, 2003b).

Nonverbal Working Memory (NWM)

The Working Memory subtests within the nonverbal domain begin by assessing the individual's ability to hold fundamental, observable objects in short term memory and progress into a rote memory block tapping task. However, towards the higher end of the subtests, the information presented becomes less concrete and more complex (Roid, 2003b).

Verbal Fluid Reasoning (VFR)

The Fluid Reasoning subtests within the verbal domain measures reasoning, absurdities and analogies. As mentioned earlier, the individual is required to sort, identify

what is absurd or impossible about verbally presented sentences and pictures, to make generalizations about the information provided (Roid, 2003b).

Verbal Knowledge (VK)

The Knowledge subtests within the verbal domain are Vocabulary. The subject is required to identify several objects and perform through picture vocabulary. As the difficulty level increases, the subject must clearly define vocabulary words. At the upper levels, performance on this subtest is influenced by schooling (Roid, 2003b).

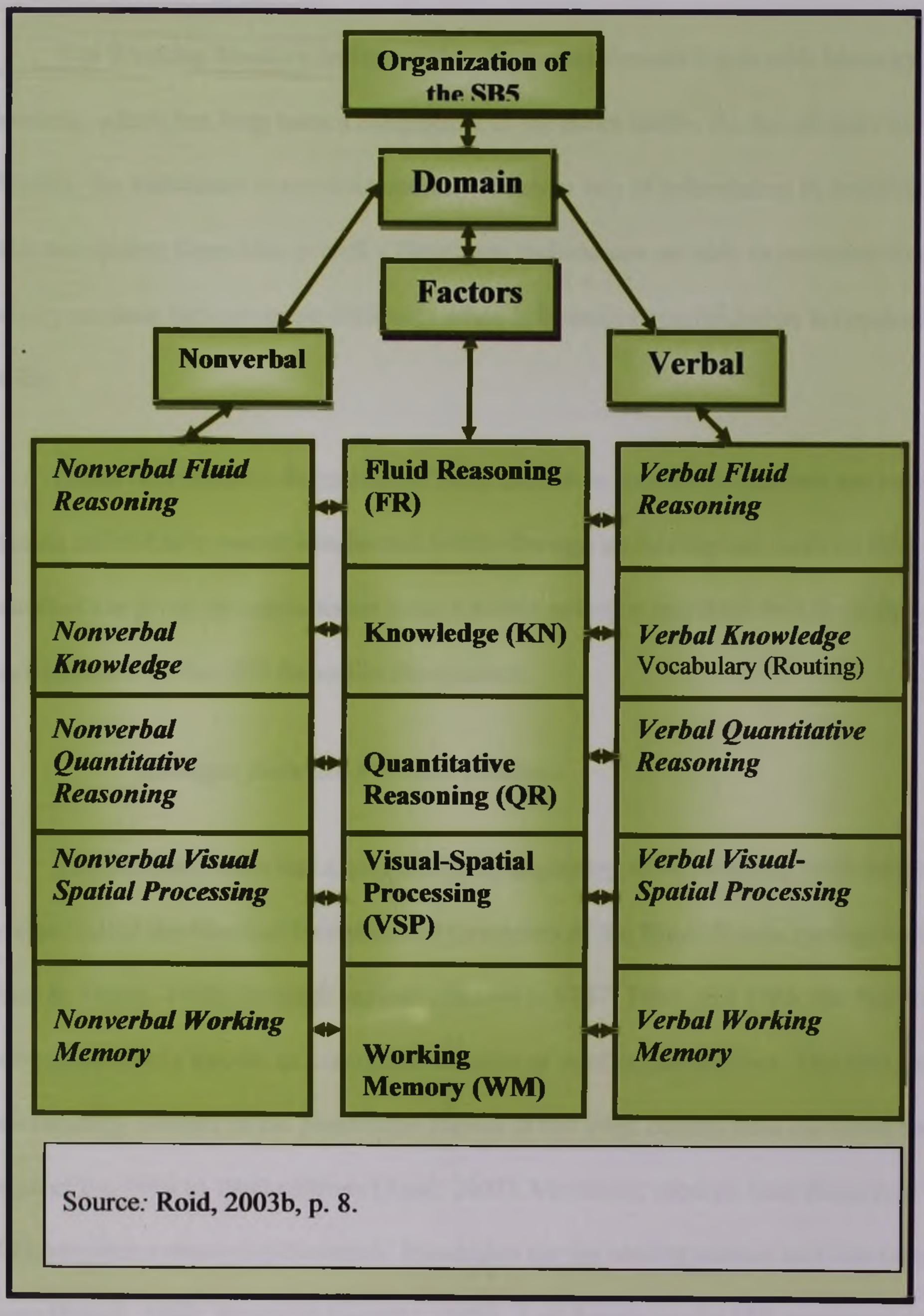
Verbal Quantitative Reasoning (VQR)

The Quantitative Reasoning subtests within this verbal domain measure an individual's ability to use a variety of mathematical skills. The subtest assesses the individual's basic addition and subtraction skills, geometric, measurement skills and to complete word problems involving multiplication at difficulty level (Roid, 2003b).

Verbal Visual Spatial Processing (VVSP)

The Visual-Spatial Processing subtests within the verbal domain assess the individual's ability to understand spatial concepts and relationships. The lower levels of the test include terms such as "ahead" and "behind," and do not rely heavily upon expressive vocabulary. However, as the task demands increase, expressive vocabulary is needed to explain the complex relationships between geographic information (Roid, 2003b).

Figure 2. Organization of the SB5.



Verbal Working Memory (VWM)

The Working Memory testlets within the verbal domain begin with Memory for Sentences, which has long been a component of the Binet scales. As the subtests increase in difficulty, the individual is required not only to retain bits of information in working memory, but to manipulate these bits as well. Oftentimes, individuals are able to complete the rote memory sections but encounter difficulty when information manipulation is required (Roid, 2003b).

Based on extensive discussion of SB5, the above mentioned subtests are basic key to judge an individual's overall intellectual ability through an intelligence scale as SB5. Thus, the author has given an emphasis on these subtests as major variables for her study as standardization of the SB5 for use in Bangladesh.

Changes from the Previous Editions.

The Stanford-Binet has a long tradition, beginning with Terman's 1916 American revision called the Stanford Revision and Extension of the Binet-Simon Intelligence Scale (Binet & Simon, 1908).Through various editions in 1937, 1960, and 1986, the Stanford-Binet has become widely known as a standard measure of intellectual abilities. The SB5 blends the use of routing subtests in the point-scale format of the 1986 edition with the functional level design of the 1916 to 1960 editions (Roid, 2003). Moreover, modern Item Response Theory (IRT) provides a strong psychometric foundation for the routing subtest and functional-levels design (Rasch, 1980; Wright & Uneacre, 1999). Test design for the SB5 employed many of the "new rules of measurement," based on IRT, recognized by psychometric experts (Embretson, 1996; Embretson & Hershberger, 1999; Reckase, 1996). These new measurement rules include methods such as calibrating items in an extensive item pool and

adaptive testing through the use of routing subtests. By adapting the test, the routing procedure of the SB5 increases the precision of measurement by tailoring the level of item difficulty to the examinee's level of cognitive functioning. Traditionally, routing has been a unique feature of the Stanford-Binet scales. Many of the familiar subtests of previous editions remain in the SB5. Examples include Picture Absurdities, Matrices, Vocabulary, and Memory for Sentences, Quantitative Reasoning, and Verbal Absurdities. The use of a hierarchical model of intelligence (with a global g factor and multiple factors at a second level in Fig 3), established in the Stanford-Binet Intelligence Scale: Fourth Edition (SB4) (Thorndike, Hagen, & Sattler, 1986) is repeated in the SB5. A few classic items, such as those in picture absurdity, have been included in the new edition to provide consistency across editions. Changes from the Fourth Edition include a general modernization of artwork and item content as well as the following enhancements (Roid, 2003).

Additional factor.

The SB5 includes five factors (Fluid Reasoning, Knowledge, Quantitative Reasoning, Visual-Spatial Processing, and Working Memory) instead of the four factors in the SB 4.

Child-friendly materials.

Responding to many user requests, the SB5 brings back many of the toys and colorful manipulative that are engaging for small children and helpful for early-childhood assessment.

Enhanced nonverbal content.

One half of the subtests in the new edition employ a nonverbal mode of testing, requiring no, or minimal, verbal responses from the examinee. Unique to the SB5, compared to other intelligence batteries, is that the Nonverbal IQ covers all five major cognitive factors.

Increased Breadth of the Scale.

New items to measure very low functioning and very high giftedness have extended the scales upward and downward to provide a wider range of assessment. For example, Object Series items were added to the lower end of Matrices to provide an exceptional floor for the routing tests.

Enhanced usefulness of the test.

The types of items, scores, and factors for the SB5 have been designed to facilitate clinical use of the SB5. The contrasts between verbal and nonverbal facets of each of the five factors, the Abbreviated and Nonverbal forms of the test, and the Working Memory subtests enhance the interpretations and applications of the test in clinical, school, and occupational settings. Based on the description of changes from earlier editions, the unique features of the SB5 are as follows (Maddox, 2003):

- Wide variety of items requiring nonverbal performance by examinee - ideal for assessing subjects with deafness or communication disorders.
- Ability to compare verbal and nonverbal performance - useful in evaluating learning disabilities.

- Greater diagnostic and clinical relevance of tasks, such as verbal and nonverbal assessment of working memory.
- Extensive high-end items, many adapted from previous Stanford-Binet editions and designed to measure the highest level of gifted performance.
- Improved low-end items for better measurement of young children, low functioning older children or adults with intellectual disability.
- Co-normed with measures of visual-motor perception and test-taking behavior.
- Enhanced artwork and manipulative that are both colorful and child-friendly.

The Standardization of (Original) 2003 Edition (SB5)

The total of ten subtests, five nonverbal and five verbal provides measures of the five CHC factors in the SB5: Fluid Reasoning, Knowledge, Quantitative Reasoning, Visual-Spatial Processing and Working Memory. Out of nearly 1000 items from the pilot and tryout phases of the project, approximately 375 items were employed in the 5th standardization edition. The final published version separated the nonverbal and verbal subtests into separate easel books whereas the longer Standardization Edition had a mixture of nonverbal and verbal subtests in each functional level of the test. Very close statistical equivalence for the two versions (longer Standardization Edition and shorter final version) was demonstrated, and no significant context or order effects were observed between the two versions (Roid, 2003).

Psychometric Properties of (Original) SB5 for Standardization

Extensive studies of reliability, validity, and fairness were conducted as part of the SB5 standardization.

Item Analysis.

The items from all Stanford-Binet editions were rated by experts in the Cattell-Horn-Carroll (CHC) theory of intellectual abilities during the first year of the development of the SB5(Carroll, 1993; Cattell, 1963; Evans, Floyd, McGrew, & Leforgee, 2001; Horn, 1994). The experts noted the CHC factor or factors being measured by each item, and all items were classified into comprehensive lists for each factor. These lists proved valuable in creating early versions of new items and new subtests. Factor analyses of Forms L and M of the Stanford-Binet Intelligence Scale (Terman & Merrill, 1937) and the Stanford-Binet Intelligence Scale: Fourth Edition (Thorndike et al., 1986) further verified the items and subtests most central to each of the factors. Extensive item analyses, including classical and item response theory methods, were conducted on SB5 items. Item analyses, subtest scaling analyses, reliability studies, and item factor analyses were conducted using pilot, tryout edition, and standardization edition studies. The final selection of items for the standardization edition involved many sources of information, item analyses, and the comparative merit of items. Following the national standardization, all items and subtests in the SB5 were again scrutinized and some items were removed for the final published edition. Some of the criteria and data employed in the final item selections were the following:

- Freedom from gender, racial/ethnic, cultural, or religious objections or differential item functioning based on expert reviews and statistical indexes of potential item bias.
- Excellent fit of the item to the one parameter logistic (Rasch) model for each of the five dimensions of the battery.
- Strong recommendations from examiners in pilot, tryout and standardization studies.

- High subtest internal-consistency and interscorer reliability and high discrimination indexes for items.
- Appropriateness of difficulty and range of difficulty for the expected age or functional-ability range.
- Positive contribution to the factor structure (five cognitive dimensions) and the total test.
- Evidence of content, criterion, and construct validity (e.g. differentiation of normative and special or exceptional examinee groups; factor structure).
- High ratings by users of previous Stanford-Binet editions, experts, and advisory-panel members.

Norm.

The sample was nationally representative and matched to percentages of the stratification variables identified in U.S. Census Bureau (2001) publications. The stratification variables were age, sex, race/ethnicity, geographic region and socioeconomic level, each of which is being defined below.

Age.

For stratified sampling purposes, 30 age groups were defined. Age was defined by subtracting the birth date from the testing date, with months of age treated as 30 days.

Sex.

Either examinees or their parents or guardians identified the sex of the examinee on the required consent form. Examiners verified sex by interview (and by markings on the SB5 Record Form) if this information was missing or unclear. An approximate 50% split between

female and male examinees were targeted at all age levels except the elderly, where census studies clearly show a larger percentage of females.

Geographic region.

The four U.S. geographic regions in the census (Northeast, Midwest, South, and West) were employed in stratifying the normative sample. The home or usual residence of the examinee denned the region, not the school or agency where testing was conducted.

Socioeconomic level.

As with numerous other published instruments in psychology and education, educational attainment was employed as the indicator of socioeconomic level. The other popular indicators of socioeconomic standing, occupation and income, were judged to be problematic. Although occupational information was collected for the SB5, it is by nature a complex description of the jobs of the parents or guardians that would then have to be categorized by various scales of occupational level—a time-consuming and fairly subjective process. (U.S. Bureau of the Census, 2000).

Reliability.

Scores obtained from tests of intellectual ability such as the SB5 must be as precise as possible, given that they are used for life-changing decisions of treatment, placement, or classification. However, the concept that all test scores have some degree of measurement error is critically important to the ethical use of tests (Turner et al., 2001). Measurement error is evaluated by examining the reliability of each test score. The reliability of a test score refers to its precision in measuring the true attributes of a person and its consistency across sets of items, multiple testing occasions and other conditions that affect score stability. Reliability for SB5 scores includes internal consistency, test-retest stability and errors of measurement. Internal-consistency reliability ranged from 0.95 to 0.98 for IQ scores and from 0.90 to 0.92 for the five factor index scores. For the 10 subtests, average reliabilities

(across age groups) ranged from 0.84 to 0.89, providing a strong basis for profile interpretation. Test-retest reliability studies were also conducted and showed the stability and consistency of SB5 scoring (Roid, 2003).

Validity.

Validity has numerous features and is established by the presentation of content-related, criterion-related and construct-related evidence. Validity is assessed by correlating measures with a criterion measure known to be valid. Evidence for content and criterion-related validity of the SB5 was conducted. Examples of validity, including the correlations with other assessment batteries were computed for standardization of original SB5. The correlations shown are quite substantial and similar in magnitude to the concurrent correlations observed for other major intelligence devices (Roid, 2003). Besides, the research related to the foundation of the five key factors of SB5 is reviewed below.

Research Related to the Stanford-Binet Intelligence Scale (SB5)

The landmark research of Carroll (1993) based on 461 factor studies of intelligence has resulted in an integrated theory of intellectual ability which was regarded as the leading research-based model of intelligence. The integration of Carroll's work with previous research has lead to the new Cattell-Horn-Carroll (CHC) theory of intellectual abilities (Flanagan, 2000; Evans et al., 2001). As a result, the selection of the CHC model allows the SB5 and its users to benefit from more than 60 years of accumulated research and clinical experience in the assessment and interpretation of intellectual abilities. Studies on the early Stanford-Binet Forms L and M showed that the CHC factors were clearly recognizable in the early editions of the Binet scales (Woodcock & McGrew, 1997), adding an even greater degree of historical and clinical meaningfulness to the CHC model. Figure 3 shows the CHC model with the five factors of the SB5 displayed in the middle row below general ability (g).

Five Factor Model.

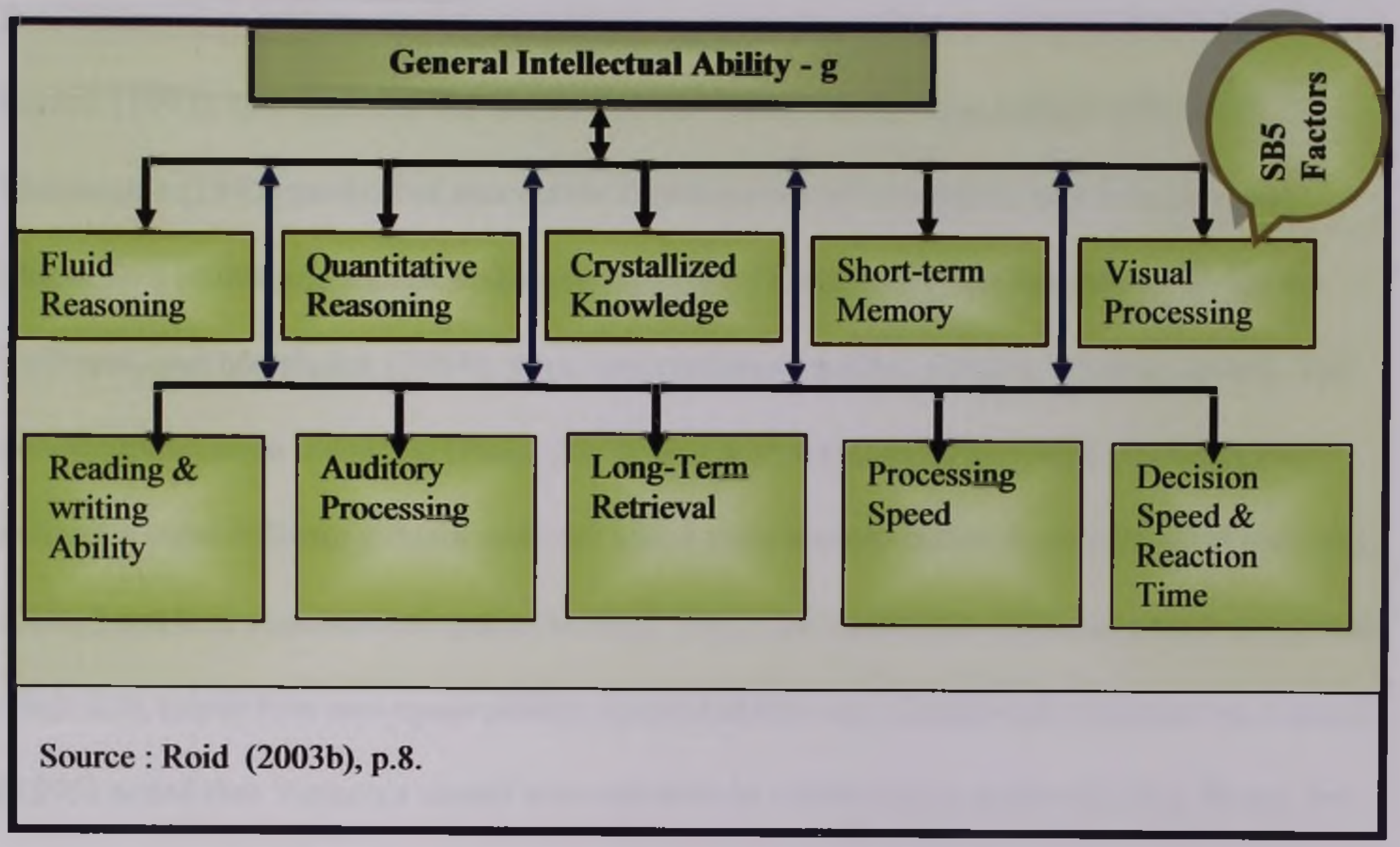
The SB5 includes five factors from the CHC model (as shown in Figure 3). The importance of these five factors emerged from extensive review of the literature on intellectual assessment and extensive discussions with experts in giftedness, special education, pre-school assessment and adult-clinical disorders.

The SB5 factor names, SB5 factor abbreviations, and CHC equivalent of the five SB5 factors are listed below (Roid, 2003).

- Fluid Reasoning (FR) (Fluid Intelligence or Gf)
- Knowledge (KN) (Crystallized Knowledge or Gc)
- Quantitative Reasoning (QR) (Quantitative Knowledge or Gq)
- Visual-Spatial Processing (VS) (Visual Processing or Gv)
- Working Memory (WM) (Short-Term Memory or Gsm)

These five factors tend to have the highest g loadings in the CHC model, especially FR, KN and QR. The five factors include those predictive (Evans et al., 2001) of school achievement (Gc, Gsm, Gq), and those generally recognized by experts in giftedness (Benbow & Lubinski, 1996 in Roid, 2003) as key elements of higher-order thinking and general reasoning ability (Gf, Gc, Gq). The Visual-Spatial Processing factor was included to provide a strong nonverbal content to the SB5 and because of its importance in identifying students with spatial talent. Also, these key factors have been identified as prominent among the abilities measured by the previous editions of the Stanford-Binet from 1916 to the present (Woodcock, & McGrew, 1997; Roid, 2003).

Figure 3 . The Structure of Cognitive Abilities.



Higher Order Factors in Studies of Intelligence and Cognition.

Hierarchical studies of intelligence and cognition originally grew out of Spearman's (1927) model of general intelligence, which he labeled g. More recent descriptions of hierarchical models appear in Carroll (1993), who proposed a three-stratum model with numerous specific factors in stratum one, eight factors in stratum two and 'g' in stratum three. Carroll's eight factors include those listed in the CHC model (Figure 3) except that he placed quantitative reasoning as part of fluid reasoning and reading and writing ability as part of crystallized knowledge. He also stated that he accepted the basic features of Spearman's concept of g and the enhancements developed by his colleague Holzinger (1936). This later Spearman-Holzinger model was similar to Carroll's hierarchical three-stratum theory, except that it included only the top two strata—the specific group factors and the g factor (Roid, 2003).

Besides, Thurstone (1938) identified seven primary mental abilities: verbal, word fluency, number facility, spatial visualization, reasoning, memory and perceptual speed. Carroll (1993) indicated that the modern three-stratum model was a direct outgrowth of Thurstone's (1947) method of successive factorization of correlation matrices at higher orders. In a multidimensional scaling reanalysis of Thurstone's data conducted by Snow, Kyllonen, and Marshalek (1984), three superordinate clusters of tests—verbal, spatial, and quantitative—were identified (Roid, 2003). In another hierarchical model, Vernon (1961) defined a superordinate g factor and two lower order factors called v:ed (verbal-educational ability) and k:m (mechanical-spatial ability). The v: ed subdivides into verbal and numerical, while k:m subdivides into space ability, manual ability and mechanical information. Carroll (1993) noted that Vernon's model was valuable in confirming a hierarchical *g* factor, but was oversimplified in claiming only two lower order factors (Roid, 2003).

Cattell (1943) developed the initial fluid and crystallized model of intelligence. Cattell considered fluid intelligence to consist of deductive and inductive reasoning and the ability to solve novel problems. Crystallized intelligence involved the processing of accumulated knowledge due to acculturation, schooling, language development and general ability to reason with stored information and methods. Horn (1965) confirmed the fluid-crystallized distinction, but added other factors now identified as visual-spatial ability, short-term memory, processing speed, and long-term retrieval (Horn & Cattell, 1966). Quantitative reasoning was identified by Horn (1989) and in the cross-battery factor analyses of Woodcock (1990).

Independent of other investigators, Gustafson (1984) proposed a three-level model of intelligence. At the highest level is *'g'* (general intelligence) and at the next level are three broad factors. These factors are labeled crystallized intelligence (dealing with verbal

information), fluid intelligence (ability to solve novel problems), and general visualization (dealing with figural information). In Gustafson's data, fluid intelligence showed an extremely high relationship to the higher order '*g*' factor, suggesting that fluid reasoning is at the core of general intelligence. At the third level are the primary factors of verbal and numerical achievement within crystallized intelligence. Also, speed of closure, figural relations, induction and memory span are found within fluid intelligence. Finally, visualization, spatial orientation and flexibility of closure are found in general visualization. Gustafson and Undheim (1992) replicated these findings with 12- and 15-year-olds. They reported a general intelligence factor with residual factors representing crystallized intelligence (read as verbal) and general visualization (read as figural or nonverbal). The consistency between Gustafson's model and the major factors in CHC theory and the SB5 are striking (Roid, 2003). The varied independent investigations of Spearman-Holzinger, Carroll, Thurstone, Cattell-Horn and Gustafson converged on the types of factors found in the CHC model. In all of these studies, prominent verbal and visual factors emerge as important along with quantitative, memory, and reasoning factors, providing indirect support for the SB5 verbal-nonverbal dichotomy and the five-factor model (Roid, 2003).

Historical Antecedents of the Non verbal-Verbal Domain of the SB5.

Alfred Binet was clearly aware that intellectual behavior could occur without language-based determinants. As his development of intelligence measures evolved, Binet (1903) became aware of the fact that intelligence may be formulated through thinking without images and thinking without words:

> The images, the interior language, and the acts are the conscious forms of the thought; they are its light; they render the thought visible to us, they reveal its details to us.... But they come only after the thought, they are its results; before the images, before the words, the thought is understood, it is performed. . . . We believe that we have established beyond any doubt, by precise observations, that there is thought without images, that there is thought without words, and that thought is formed by an intellectual feeling (Binet & Simon, 1908, pp. 338—339).

Terman and Merrill (1937) sought to increase the number of nonverbal procedures at the lower levels of the Stanford-Binet, motivated by concerns with the verbal nature of the lower items. Several subsequent attempts have been made to create nonverbal scales for the Stanford-Binet. In the first attempt, McNemar (1942) created two 20-item parallel nonverbal intelligence scales from items in the Stanford-Binet Forms L and M, although he noted some limitations: *"Since the directions for these items are mainly verbal rather than pantomime, it follows that some understanding of language is involved and consequently that the items are not to be regarded as purely non-verbal"* Roid (2003). In a recent attempt to develop a nonverbal short form, Glaub and Kamphaus (1991) selected subtests from the Fourth Edition for use with children who are hard-of-hearing, speech and language impaired, and limited-English-proficient children. This nonverbal short form consisted of Bead Memory, Pattern Analysis, Copying, and Memory for Objects and Matrices.

More recent advances in test development suggest that multidimensional tests may be developed that are largely nonverbal (Roid & Miller, 1997; Roid & Holadyna, 1982). These tests usually involve pantomime administration, content and structural studies regarding the degree to which item performance is mediated through verbal or nonverbal means and varied

item response modes (e.g., pointing to a stimulus book, placing cards in an appropriate arrangement, building three-dimensional constructions). The SB5 development team (Roid et.al, 2003) studied these nonverbal innovations and other previous methods of verbal and nonverbal assessment. Accordingly, it was possible to construct verbal and nonverbal Stanford-Binet scales for each of the five factors assessed in the SB5.

Cross –Cultural Assessment and Standardization Process – An Overview

During the past several decades, the unique challenges of cross-cultural assessment and counseling have attracted considerable attention. Cross-cultural assessment has become a sensitive issue due to specific concerns regarding the use of standardized tests across cultures (Chang, 2008). Before selecting an assessment instrument for use in counseling or research, counselors and researchers are trained to verify that the test is appropriate for use with their population. In order to assess overall performance, most psychological tests employ a standardization process. It allows the test developer to create a normal distribution which can be used for comparison of any specific future test score. The term standardization refers to the process of determining established norms and procedure for a test to act as a standard reference point for future test results. The criteria of standardization for any psychological test are as referred to as item analysis, norm development, reliability and validity (Anastasi & Urbina, 1997). This depicts that the investigation of validity, reliability and appropriate norm groups to which the population is to be compared.

Hence, the researcher has followed and established standardization process based on literature review of SB5 for her study. It allows the test developer to create a normal distribution which can be used for comparison of any specific future test score. The standardization process includes following steps.

Item Analysis.

Item analysis is the process of collecting, summarizing and using information from students' responses to assess the quality of test items. The analysis depicts the effectiveness of the items in a given test that discriminate between students with higher and lower scores in the ability measured. Presence or absence of faults logically affects the values of discrimination. Items that discriminate poorly indicate adaptation and modifications. Difficulty Index (P) and Discrimination Index (D) are two parameters which help to evaluate the standard of test items used in an assessment. (Mitra, Nagaraja, Ponnudurai & Judson, 2009).

Adaptation

A psychological test would be effectively standardized for using in other language and culture when the items of the test are being well adapted. With this view, standardization and adaptation of psychological tests for cross-cultural assessment is becoming prominent worldwide at an increasing pace. For years, psychological and educational tests have been translated for use in different languages (Geisinger, 2003cited in Matthews, 2003). The contribution that adaptation brings to the process is the flexibility to make an altered instrument not only linguistically appropriate but also culturally fit with an intended targeted population. (Hambleton & Bollwark, 1991cited in Matthews, 2003). Moreover, adaptation and translation of tests are not limited to the areas of academic testing. In particular, researchers in the psychological arena are also using adapted tests to assess intelligence, aptitude, personality etc (Chang & Myers, 2003cited in Matthews, 2003). Adapted tests are usually a mixture of three types of items: newly developed items, translated items and adapted items (Church, 2001).The variation in adaptation depends on the function and the

purpose of the test being used (Hambleton & Bollwark, 1991 in Matthews, 2003). Verbal tests, for example, may require more new item development and item adaptation whereas a test of mathematical reasoning may consist primarily of item translation and corresponding adjustments to instructions.(Matthews,2003). A safe rule of thumb to translate or adapt an item is to ensure whether the proposed target item reflects the spirit of the original item (Allalouf & Chang, 1999; Sireci, 1998 cited in Matthews, 2003). Moreover, adapting an existing instrument instead of developing a new one has remarkable benefit. By adapting a test, the researcher is able to compare the cross-cultural studies at both the national and international level. For any test developers and users, adaptations also conserve time and expenses (Hambleton, 1994). Test adaptation can lead to increased fairness in assessment by allowing individuals to be assessed in the language of their choice (Hambleton, Merenda & Spielberger, 2005). On the other hand, along with newly developed tests, in other context, tests are also being translated and adapted in different cultures. Test adaptation is exclusively necessary for language and cultural differences (Reckase 1989). The most significant recognition for any test adaptation in any country is the provision of guidelines from the International Test Commission (ITC). ITC recommends a guideline for Test Translation and Adaptation is: "Test developers/publishers should ensure that the adaptation process takes full account of linguistic and cultural differences in the intended populations." (Guideline D1, ITC, 2001; Hambleton, 2005). This guideline can act as a benchmark for any country in translating and adaptation of psychological tests. Examples of psychological assessments available in various languages and culture, including intelligence and general ability tests, Stanford-Binet Intelligence Scale Fifth Edition (SB5), the Wechsler Intelligence Scale for Children- Revised (WISC-R) , the Peabody Picture Vocabulary Test, the Bateria Woodcock-Munoz, the Children's Hope Scale, the Sixteen Personality Factor (16PF) Questionnaire, Miller's Analogy Test (MAT) and the Wonderlic Personnel Inventory

are adapted instruments. (Matthews, 2003). These adapted tests are widely used over the world, especially in developing country for increasing fairness and usability in assessment.

Norm.

Norms are not standards of performance, but serve as a frame of reference for test score interpretation. Norm groups can range in size from a few hundred to a hundred thousand people. The more people are used in norm group, the closer the approximation to a normal distribution. The standardization sample is also referred to as the norm group. Generally for standardization, the samples are representative and matched to percentages of the stratification variables such as age, sex, race/ethnicity, geographic region and socioeconomic level (described in chapter two) (Overton, 1992).

Reliability.

Reliability in assessment refers to the confidence that can be placed in an instrument to yield the same score for the same student if the test is administered more than once. Besides, it considers the degree to which a skill or trait is measured consistently across items of a test. Since educators use assessment as a basis for educational intervention and placement decisions, understanding of reliability aids educators in determining the accuracy and dependability of an instrument (Overton, 1992). The reliability of a test score refers to its precision in measuring the true attributes of a person and its consistency across sets of items, multiple testing occasions and other conditions that affect score stability (Rousson, Gasser, & Seifer, 2002).

Validity.

The accumulation of evidence for the validity of test scores and their interpretation is a complex effort. Validity of test scores depends on the proper administration of the test by an experienced examiner and proper recognition of the unique characteristics of the individual examinee (Matarazzo, 1990). Technically, a test is neither valid nor invalid by itself, but instead, the uses and interpretations of test scores are valid or invalid based on accumulated evidence (Turner, DeMers, Fox, & Reed, 2001).

Thus, validity and reliability take an additional dimension in cross-cultural testing as do the question of the appropriate norm group. The instrument must be validly adapted, the test items must have conceptual and linguistic equivalence and the test items must be bias free (Domino, 2000). As stated earlier, the International Test Commission (ITC) has provided guidelines for translating and adapting tests in 1992. It further highlights administration and interpretation of tests to improve the accuracy and compile evidence on the equivalence between the different language versions (Guideline D1., ITC, 2001; Hambleton, 2005). There is considerable evidence indicating that the need for multilanguage versions of intelligence, achievement, aptitude and personality tests are growing. These adapted tests would then be appropriate if further research and cross-cultural comparative studies are being carried out (Hambleton, 1994).

Hambleton (1994; 1993) also emphasized on the reliability of the adapted version through back translation. Thus, the test is translated into the target language and then it is re-translated back to the source language. He also point out that once the process is complete, the final back-translated version is compared to the original version. In many ways, the process of test adaptation mimics that of new test development. The decision to engage in this

process is driven by the general need to produce an equivalent or comparable instrument for use in a target language and /or with a target culture population (Al-Ansari & Bella,1997; Chang,1999 cited in Mathews, 2003). When utilizing the test in a new cultural group, it is not quite as simple as directly translating the test, administering it and then comparing the results for its validity (Vijer & Hambleton, 1996).

The translation and adaptation of tests in more than one language and culture is a traditional phenomenon. For example, the Binet-Simon Intelligence Test was translated from French to English in 1911 (Hambleton & Bollwark, 1991cited in Matthews, 2003). Within just few years, the same test had been translated into seven languages. Hambleton (1993) asserts that researchers and test developers have been interested in translating psychological tests and attitude scales since the beginning of standardized testing. Hambleton (2005) makes the point that the need for test adaptation is not limited to international comparative assessments. He also asserts that translation is but a part of the adaptation process and that adaptation of scoring procedures, instructions to examinees and related questionnaires are also required (Matthews, 2003).

Interpretation of Test Scores: Now and Then

In the early decades of intelligence testing, intelligence test scores were expressed as a true quotient, hence the term IQ or intelligence quotient. An IQ was defined as a ratio of the examinees mental age to the examinees chronological age which was then multiplied by 100 to eliminate dealing with fractional scores [(MA/CA) X 100]. This form of calculation for an IQ has serious psychometric and related measurement problems and has been abandoned for decades although its presentation continues to be common in many introductory psychology and education textbooks. In the early 2000s, IQs are calculated in the form of age corrected

deviation scaled scores. These are formal transformations of raw scores (i.e., number of points obtained or items answered correctly) into a standard score format that incorporates the use of the mean and the standard deviation of the raw scores at predetermined age intervals so that the IQ given by the test has the same percentile ranking at each age level, which is not true of the old ratio style IQ. Further, IQ score is a necessarily incomplete reflection of intelligence. It is far from perfect as an index of a person's total intellectual ability and is not useful in identifying specific talents. Scores from intelligence tests are interpreted properly only when the standardized instructions for administering and scoring the test have been followed rigidly. Deviations from standardized administration and scoring cause the scores to move up or down for an individual examinee inappropriately and in ways that are unpredictable, rendering the scores uninterruptable (Lee, Reynolds, & Willson, 2003). Intelligence test scores are viewed by test interpreter as reflecting innate potential but clearly that is not the case. While innate ability contributes to intelligence test performance, many other variables contribute to performance on ability measures as well. Intelligence as measured on such tests as described here is a summative construct at any given point that is a reflection not only of a person's innate potential but the interaction of this potential with the entire life experiences of the individual (Reynolds, Livingston, & Willson, 2006).

On the contrary, the observed increase in average IQ scores of 3 points per decade has been reported as proof that intelligence is not stable but is flexible with regards to environmental influences (Flynn, 2007 & Collingwood, 2008). Besides, several studies describe that children today score higher on an old IQ test than children the same age did who took the same test decades earlier. A possible explanation for the increase in IQ scores is that children today mature sooner, both physically and mentally, than children did decades ago (Flynn, 1994). Today's children score higher, not because their real intelligence has

increased, but because their brains are more mature. There is considerable evidence that children today mature earlier. A number of reasons have been given for the earlier maturation of children. Explanations have included hereditary and diet factors, increases in obesity and body weight, chemicals acting as endocrine disrupters, and the sexualization of children by the media (Roberts & Cox, 2005). Average scores on intelligence tests are rising substantially and consistently, all over the world. These gains have been going on for the better part of a century—essentially ever since tests were invented. The rate of gain on standard broad-spectrum IQ tests amounts to three IQ points per decade, and it is even higher on certain specialized measures (Ceci, 1996).

Keeping pace with Binet's intention and cautioned (mentioned earlier) for test score interpretation and Roid's scoring procedure, the study focused only age graded norms for its scoring interpretation. The present study has taken into consideration the above criteria while going through the process of standardizing the Stanford Binet Intelligence Scale Fifth Edition for use in Bangladesh.

International Perspectives - An Overview

Since thousands of years for interest in intelligence, assessment techniques were commonly used in China for civil service testing over 4000 years ago. In the Western world, Sir Francis Galton was the first to state that there were four important ideas about the assessment of intelligence. Galton believed differences in intelligence were quantifiable in terms of degrees of intelligence, differences among individuals formed a bell shaped curve, intelligence could be measured by objective tests and these tests could be correlated (Gregory,2007).

The Binet-Simon scales, including the 1911 revision, which extended through adulthood, were welcomed in Europe and the United States and their translation and adaptation had almost immediately begun. Among several investigators, Town directly translated the Binet - Simon scale into English, in 1913 early revisions and adaptations were developed by Bobertag in Germany, Johnston and also Winch in England and Goddard, Kuhlman, Wallin, Terman and Yerkes adapted in the United States. (Pintner & Patterson, 1925).

The Stanford-Binet Intelligence Test was translated by Leuckert in German language. Amelang (1967) made a comparative study with Hamburg WISC and the Stanford-Binet Intelligence Leuckert Scale. Presentation of statistical results include tests of normal distribution of scores, means, standard deviation and range of IQ, interest correlation and the mean IQs of professional categories. In Italy, Bazzo and his associates (1966) worked with the Italian revision of the Stanford-Binet Scale Form L—M, 1960. They also discussed the third revision of the Stanford Binet Scale by Terman and Merrill in 1960 and reported minor modifications adopted for use with Italian children as a result of research conducted in Genoa (Huq, 1992).

Modern psychology is moving east with increasing rapidity in Asia especially in Japan, almost as early as it did in United States. The subject was introduced into a Japanese University around 1890. It made an appearance in China before 1920, and in some of the colonized countries, such as India and Vietnam before World War II, somewhat later in the Australian and New Zealand colonies, where the intellectual life and educational institutions of the new population were little influenced by the original inhabitants.

India had a rich treasure house of psychological knowledge along with the ancient philosophical and religious texts. In 1921 the Bureau of Education and Research of the Government of India devised a series of tests based on Stanford-Binet Intelligence Scale. Later in 1952, C. Herbert Rice devised a Hindustani Binet Scale. Further, the Bombay-Karnataka revision of the Stanford-Binet by Kamat (1955) and the Lady Willingdon Training College Revision of Binet in Tamil and Telegu turned out to be a sustained experience on the adaptation of intelligence tests. Later Kulshrestha (1971) adapted the Stanford-Binet Intelligence Scale form L-M in Hindi (Huq, 1992). Besides, Samejima Fumik developed an absolute scale in Japan on basis of the raw scores of the Suzuki-Binet Intelligence Scale (1958). In 1910, Henry Tasman Lovel was influenced by Binet's work on mental testing. Later on, Philips, the principal of Sydney teachers college in 1925 published a Sydney revision of the Binet tests (Huq, 1992). The Stanford- Binet Scale of Intelligence (Form L) was adapted and standardized for Turkish blind children between 1982-84. The findings proved some evidence that the Binet type items at the younger ages at least, holds up well cross-culturally and with the group of handicapped children.

In 1998, a study by Rodriguez, Treacy, Sowerby & Murphy evaluated the applicability of Australian adaptations of WISC- Third Edition and SB4 for Dunedin children. The research suggested that New Zealanders obtain scores at or above those of the US normative sample. The findings of the research stated that the Dunedin children obtained means comparable to the US norms at the significant level of 0.01. In a similar way several research studies were conducted on the Stanford Binet Scale which has also been reviewed in the light of literature. Likewise, in his review of SB5 in 2003, Kirk A. Becker stated and named an article "History of the Stanford-Binet intelligence scales: Content and

psychometrics." He noted the SB5 as one of the first examples of an adaptive-type intelligence scale and highly recommended the practical utilization of the test.

A comparative study between fifth edition and the fourth edition of the Stanford Binet was conducted by Kush in 2005. The findings revealed that there was significant correlation measured through criterion-related validity with Full Scale scores averaging 0.90. The Fifth Edition average Full Scale score was 107.9 and the Fourth Edition average Mean Composite score was 111.4. These results are found to be consistent with the Flynn Effect. In 2005, Danielle Chase and his colleagues reviewed a literature on "Underlying Factor Structures of the Stanford-Binet Intelligence Scales – Fifth Edition". The review focused on the exploratory factors while constructing a new test instrument. According to the findings , the Exploratory Factor Analysis (EFA) of the SB5 data indicated that a different factor structure underlies the measure. This information is very useful to clinicians and may guide interpretation of the SB5 in practice. These findings also confirmed the necessity of EFAs when constructing a new test device. Douglas and his associates (2006) conducted a study on "Males have greater g: Sex differences in general mental ability from 100,000 17- to 18-year-olds on the Scholastic Assessment Test". The study depicted that there are no sex differences in overall general intelligence. Their findings of "no sex difference in intelligence" have since been replicated many times on other standardization samples with other test devices. However, according to their opinion, males are often observed to average higher scores on some tests of spatial ability, mathematical reasoning and targeting, while females are often found to average higher on some tests of memory, verbal ability and motor coordination within personal space.

Similarly, in 2006, Mark Pomplun and Michael Custer conducted a study to "examine the validity of the measures of verbal and nonverbal working memory on the Stanford-Binet

Fifth Edition (SB5). The validity evidence included Rasch-based, criterion-referenced item mapping, correlations with other clinical measures of memory and prediction of reading and mathematics scores. The higher correlations of the SB5 verbal and nonverbal working memory subtests with other measures of verbal and nonverbal memory, respectively. Besides, the lower correlations with nonverbal and verbal memory measures, respectively and clearly show convergent and divergent validity. The higher correlations between SB5 verbal working memory and reading skills and between SB5 nonverbal memory and mathematics skills are consistent with past research.

A study on "The predictive ability of the visual-spatial measures of the Stanford-Binet intelligence scales, Fifth Edition and the Wechsler Intelligence Scale for Children-Fourth Edition Visual-spatial processing and mathematics achievement" by Eldon Clifford in 2008 suggests that processing components play a critical role in academic tasks such as reading, writing and mathematics. Furthermore, the research states that visual-spatial processing is related to mathematics achievement. Research also supports both instruments have subtests that measure visual-spatial processing. The result found that both measures could significantly predict mathematical achievement.

In 2010, Kevin McGrew conducted a research on "CHC theory of intelligence and its impact on contemporary intelligence test batteries". According to the study, the author explained the CHC theory should be considered one of the major landmark psychometric bases for the assessment of human cognitive performance. In this research, the author also claimed that the CHC is the framework to eventually better describe and explain human cognitive performance. Therefore, Lincoln, (2010) computed the test- retest reliability of the Stanford-Binet scale and the constancy of intelligence quotients on thirty samples aged 6 and 7 years. The findings revealed that the mental age changes ranged from 0 to 6 months

with a median of 2.4 months. The change in IQ varied from 0 to 8 points, with a median of 3.4 points. In 1 out of 36 tests the performance was absolutely consistent. In spite of the fact that the correlation between the first and second examinations was 0.95, in 10% of the cases there was a change of 8 points in IQ. In another study, Jacqueline (2010) critically reviewed the Stanford-Binet Intelligence Scale using split-half method, test-retest reliability and interscorer agreement. The coefficients for the subtests were determined using the split-half method. The scores were rectified with the Spearman-Brown formula. Nonverbal subtests reliability coefficients averaged between 0.85 and 0.89. Verbal subtests reliability coefficients were strong and fell between 0.84 and 0.89. A review of the Stanford-Binet Intelligence Scales, Fifth Edition by Johnson & D'Amato (2006) described that the test development process followed the standards for educational and psychological testing (AERA, APA, & NCME, 1999) and resulted in a well-designed, technically sound instrument that follows in the footsteps of earlier editions of the Binet scales but also integrated new research on intelligence into the measure. In researcher's opinion, the addition of the new subtest as working memory would be an effective strategy in assessing those with learning problems. Additionally, she emphasized on nonverbal intelligence that would be useful in the assessment of a variety of clients in our changing world.

In view of the fact that being revised in 2003 the SB5 is going through a serles of research studies for the establishment of its acceptance worldwide. In this context, Williams (2010) analyzed the internal construct validity of the SB5 using an independent sample of high-functioning students. Findings indicated that a hierarchical, five-factor, post-hoc model provided the best fit to the data. Generally, implications for school psychologists include a better understanding of the factor structure of the SB5, especially as it relates to high-

achieving children. However, a research significantly related to the present study compared the relationship between the SB5 and the WJ-III ACH with participants aged 6 to 19 years.

The results found correlations in the range of 0.50 to 0.84 using the SB5 factor index and IQ and WJ-III ACH scores (Roid, 2003 in Krystal Campbell, 2006). Carson (2004) conducted a study on the impact of IQ in classroom learning with SB5. The results suggested those both nonverbal and verbal domains are relatively stronger to facilitate learning activities in the classroom environment. The study also revealed a range of interest to educators, practitioners and psychologists in enhancing their teaching learning strategies. In another similar study, Askarian, M, Ali, A.G., Kambiz, K., & Hassan, P.S (2011) computed the diagnostic validity for new edition of Tehran-Stanford-Binet Intelligence Scale in order to identify the children with learning disabilities. In this psychometric study with classical approach, the statistical society includes all students with learning disabilities related to the primary and secondary school in Tehran provinces. The study sample size was equal with 252 students with learning disabilities who were selected based on purposive sampling. The results showed that this scale had the good diagnostic validity and desirable potential to identify students with learning disabilities. So this scale as a valid tool for identifying students with learning disabilities can be used.

A significant number of studies on adaptation and standardization of different individual intelligence tests have been done worldwide. For the relevance of the present study, the literature reviews were focused on the various editions of Stanford Binet Intelligence Scale. Till date few countries have taken the initiative to standardize the scale in their own culture which the author has pointed in the text. Besides, few other countries have also conducted several researches on different areas of intelligence with the different editions of Binet scale.

Bangladesh Perspectives

Innovations in the arena of education have been possible by unveiling the hidden potentialities of individuals by measuring their intellectual abilities. Developed countries since long have proved to overcome these limitations by making progress in educational assessment. Hence, in Bangladesh the significance of such pragmatic thought has to be emphasized. Based on the above need, it is mentioned earlier that a number of professional in the area of test development had taken initiatives for adapting and standardizing intelligence scales at the end of nineteenth century.

Abdus Sobhan (1962) was concerned with determining the predictive validity of the translated version of the "verbal reasoning" and "numerical ability" subtests of the "Differential Aptitude Test"(DAT). The findings stated that the coefficient of correlation were 0.17, 0.32 and 0.19 respectively. Accordingly, Matinuddin et al. (1964:1963:1962) had done a similar study to determine the predictive validity of the translated and adapted version of the "Differential Aptitude Test" (DAT). The coefficient of correlation between Verbal Reasoning score under study and the total marks obtained in school subjects ranged from 0.49 to 0.79. They concluded that the adapted version had shown significant improvement over those obtained on the original version of the test. An Analogies Test of Verbal Reasoning (ATVR) was developed by Gulam Mustafa in 1965. The instrument in the preliminary stage proved to be a good instrument for measuring the scholastic aptitude of the higher secondary and the college students of our country. Though these studies are not directly related to standardization of Stanford Binet but indicate the significance of assessment in the field of psychological testing.

In 1963, Islam administered Intelligence Test Scale for item analysis and selection of valid test items for the grade levels of class IX through graduation. In this study, the validity and reliability coefficient of the test was 0.40 and 0.85 respectively. Haque in 1967 developed an Individual Performance Scale of Intelligence for Children of East Pakistan for ages 5 to 11years. The scale was substantially found valid and reliable and could be used for measurement of intelligence. These studies were thoroughly reviewed from the study of Huq, 1992. In 1980, Huq conducted a study for the standardization of Wechsler Intelligence Scale for Children - Revised (WISC-R) for use in Bangladesh. The correlation between the performance scale and annual examination of the school ranged from 0.44 to 0.75. As for the full scale the correlations were between 0.06 to 0.69. The reliability coefficient was found out by the split -half technique. IQ norms were developed for Dhaka city based on the scores obtained from the standardized sample. Obviously, this was the first initiative by Huq for the standardization of intelligence scale used in Bangladesh. Then, Ferial, in 1980, standardized the Denver Developmental Screening Test (DDST) in Bangladesh for children from birth to 6 years. Based on the findings, developmental norms from 1 month to 6 years were constructed for use in Bangladesh.

A study on item analysis of Stanford-Binet Intelligence Scale (Fourth Edition) was conducted by Sharmin Huq in 1989 to determine the item difficulty and discriminating power for reorganizing and rearranging the item of the subtest for further standardization process. Based on the results, no items were eliminated or substituted. The difficulty index ranged from 0.99 to 0.004 in the urban sample while it ranged from 0.09 to 0 .55 in the rural sample. In 1992, again Sharmin Huq conducted a research on the "Determination of Reliability and Validity of Stanford-Binet Intelligence Scale (Fourth Edition) and the Construction of Norms for use in Bangladesh for children aged from 3 to 9 years". The study included only the

nonverbal subtests. Contrast validity was computed and the result indicated significant difference between the two groups at 0.001 levels. And the correlation of test-retest reliability was found 0.81. Separate norms were calculated for the urban and rural samples of Bangladesh. Since then, the era of research as well as work in intellectual assessment with Stanford Binet in identifying, educational decision making and placement for children with special needs had been started in our country.

The present study can be regarded as the renovation work in relation to the previous study where only the nonverbal subtests were standardized in the earlier edition. Moreover, it is obvious that to standardize a test, the task is comprehensive as well as complex and must have a valid time frame for its usability over time. To add further, it is also unfortunate that between 1992 and till date there has been no initiative either to develop or standardize any intelligence test. In addition, very few numbers of non-standardized, informal or teacher made tests might be developed by people all through the year for their own interest and necessity of assessing academic performance of students. But the researcher pointed out only those studies that are relevant to the psychometric testing.

Review of International Commitments, National Policies and Legislation with regards to the Education of Children with Special Needs

No singular definition of the term "special needs" exists, although the term is widely used within the disaster services and emergency management industry to address people with disabilities. However, the term "special needs" is currently under debate in the disability, healthcare, and emergency management communities. "Special needs" can be narrowly defined as a broad and overarching concept. Currently, there is no federally mandated or suggested definition being provided to states and localities. In fact, federal agencies use

different definitions in addressing special needs populations and their own missions. While the meaning of the term "special needs" depends on the community, there are some terms that have legal implications and must be considered for evacuation planning.

On the other hand, a disability is an umbrella term, covering impairments, activity limitations, and participation restrictions. Impairment is a problem in body function or structure; an *activity limitation* is a difficulty encountered by an individual in executing a task or action; while a *participation restriction* is a problem experienced by an individual in involvement in life situations. Thus disability is a complex phenomenon, reflecting an interaction between features of a person's body and features of the society in which he or she lives (World Health Organization, 2012).

The Universal Declaration of Human Rights (1948)

On December 10th, 1948 the General Assembly of the United Nations adopted and proclaimed the Universal Declaration of Human Rights. The issue of Education is particularly mentioned in Article 26 & 27 in this document:

I. Article 26

(1) Everyone has the right to education. Education shall be free, at least in the elementary and fundamental stages. Elementary education shall be compulsory.

(2) Education shall be directed to the full development of the human personality and to the strengthening of respect for human rights and fundamental freedoms. It shall promote understanding, tolerance and friendship among all nations, racial or religious groups, and shall further the activities of the United Nations for the maintenance of peace.

(3) Parents have a prior right to choose the kind of education that shall be given to their children.

II. Article 27

(1) Everyone has the right freely to participate in the cultural life of the community, to enjoy the arts and to share in scientific advancement and its benefits.

UN Convention on the Rights of the Child (CRC) - 1989

Apart from two countries, this convention has been ratified by all the member states of United Nations. The four principles of CRC (Non-discrimination: Article # 2, Best Interest of the Child: Article # 3, Survival & development: Article # 6 and Participation: Article # 12) applies to children with disabilities also. Article # 28 of CRC insists that all children have the right to education on the basis of equal opportunity & Article # 29 emphasizes that the education of children shall be directed to: the development of a child's personality, talents and mental and physical abilities to their fullest potential; the development of respect for human rights and fundamental freedom… ; parents, own cultural identity, language and values including national values… and the participation of the child for a responsible life in a free society…. Etc.

Education for All (EFA): Jomtien (1990)

The basic idea of inclusion can also be found in the Jomtien Declaration. Here, Education for All (EFA) emphasizes the inherent right of every child to a full cycle of primary education, and commitment to a child-centered pedagogy, where individual differences are accepted as a challenge, and not as a problem. The Jomtien Declaration also emphasizes the need for improvement in the quality of primary education and teacher

education, recognizing and respecting the wide diversity of needs and patterns of development among primary school children.

Salamanca Declaration (1994) World Conference on Special Needs Education –

This international declaration states "Schools should accommodate all children's conditions". Inclusive education was adopted at the World Conference on Special Needs Education (SNE) as a principle in addressing the learning needs of various disadvantaged, marginalized and excluded groups. This includes children with disabilities and gifted children, street and working children, children from ethnic minorities, refugee children and other marginalized or disadvantaged children. In this context "special education needs" refers to all children that experience barriers in equal access and equal participation in education. SNE, since the Salamanca Declaration, is viewed as an integral part of all Education for All (EFA) discussions.

Standard Rules on the Equalization of Opportunities for Persons with Disabilities (1993)

The UN "Standard Rules on the Equalization of opportunities for persons with disabilities" comprised 22 Rules. The Rule 6. *Education: 'States should recognize the principle of equal primary, secondary and tertiary educational opportunities for children, youth, and adults with disabilities, in integrated settings.*

Dakar Framework (2000)

The need for inclusive education has been repeated in the notes on the Dakar Framework for Action, which mentions "...In order to attract and retain children from marginalized and excluded groups, education systems should respond flexibly. ...Education systems must be inclusive, actively seeking out children who are enrolled and responding in a

flexible way to the circumstances and needs of all learners". The achievements 10 years on since EFA have been assessed and analyzed. The Jomtien goals have not been reached and some of them were taken on board again in Dakar, extending the time for achieving the goals.

E-9 Declaration (2000)

The declaration on EFA was agreed upon during the fourth summit of the nine high population countries (which includes Bangladesh) in February 2000, and also highlights as one of the main goals that "all children with special needs will be integrated in mainstream schools."

Children with Disabilities in No Child Left Behind

The **No Child Left Behind Act of 2001 (NCLB)** is a United States Act of Congress concerning the education of children in public schools. Several key pieces of legislation over the past three decades have contributed to the evolution of the assessment process for young children with special needs. Specifically, the Education for All Handicapped Children Act amendments (P.L. 99–457, 1986), later renamed the Individuals with Disabilities Education Act (IDEA, P.L. 102–119, 1998), the 1997 version of IDEA (P.L. 105–17, 1997–1998), the 2001 Elementary and Secondary Education Act (No Child Left Behind, P.L. 107–110), and the most recently authorized 2004 version of IDEA (Individuals with Disabilities Education Improvement Act, P.L. 108–446) have all provided critical guidelines for the identification, assessment, and treatment of young children with special needs. While, initially, the focus of legislation was to merely identify children in need of early intervention services, there has been an increased emphasis in the most recent legislation (IDEA 2004; NCLB) on looking ahead to school-based services. No Child Left Behind requires all government-run schools receiving federal funding to administer a state-wide standardized test annually to all students.

This means that all students take the same test under the same conditions. The students' scores determine whether the school has taught the students well. The No Child Left Behind Act (NCLB) includes incentives to reward schools showing progress for students with disabilities and other measures to fix or provide students with alternative options than schools not meeting the needs of the disabled population. The law is written so that the scores of students with IEPs and 504 plans are counted just as other students' scores are counted. Schools have argued against having disabled populations involved in their AYP measurements because they claim that there are too many variables involved.

Biwako Millennium Framework for Action (Asia-Pacific Decade of Disabled Persons 2003- 2012)

Early Detection, Early Intervention and Education

Less than 10% of children and youth with disabilities have access to any form of education compared with an enrolment rate of over 70% for non-disabled children and youth in primary education in the Asian and Pacific region. This exclusion from education for children and youth with disabilities results in exclusion from opportunity for further personal, social and vocational development. Three targets are set for these problems:

- Children with disabilities will be an integral part of the population targeted by Millennium.
- Development Goal Target 3, which is to ensure that, by 2015, children everywhere, boys and girls alike, will be able to complete a full course of primary schooling.
- By 2010, at least 75% of children and youth with disabilities will be able to complete a full course of primary schooling.
- By 2012, all infants and young children (0 - 4 years) will have access to and receive community-based early intervention services.

Action for this area includes adequate legislation for inclusive education and national data collection on children with disabilities (0 - 16 years).

Millennium Development Goal (MDG)

Among the Eight Goals, Goal Two have focused specially on Education.

Goal 2: Achieve Universal Primary Education

Target: *Ensure that, by 2015, children everywhere, boys and girls alike, will be able to complete a full course of primary schooling.*

DPI Position Paper on Inclusive Education

'Disabled People International (DPI) believes that education should be accessible to all who desire to be educated, no matter their ability; disabled people should have the option to be integrated with the general school population, rather than being socially and educationally isolated from the mainstream without any choice in the matter. Students who are deaf, blind or deaf-blind may be educated in their own groups to facilitate their learning, but must be integrated into all aspects of society'.

The National Policy on Education in Bangladesh

Disability Welfare Act- 2001

The National Disability Welfare Act-2001of Bangladesh emphasized: establishing specialized education institutions in order to cater for the special needs of different types of disabled children, designing and developing specialized curriculum and production of text books; creating opportunities for free education to all children with disabilities below 18 years of age and provide them with books and equipment free of cost or at low-cost; endeavor to create opportunities for integration of students with disabilities in the usual classroom setting of regular normal schools wherever possible; arranging training for the teachers and

other employees working with the disabled and to arrange easy transport facilities for attending school.

The National Literacy Goal of Bangladesh

The National Literacy Goal of Bangladesh is to ensure 100% literacy rate by the year 2015. If this target is to be achieved, the education needs of children with disabilities cannot be ignored. But there is no specific mention about inclusive education or any specific intervention to address the issues of educating children with disabilities.

National Education Policy (2000)

Chapter 18: *Special Education, Health and Physical Education, Scout and Girls Guide*

Special Education

The children unable to fulfill requirements of their daily life due to physical and mental problems need special education, competent remedial measures, special care and nursing. The deaf, blind, physically handicapped, mentally handicapped and the epileptics fall within the purview of special children. In accordance with the degree of disability, they are termed as mildly, moderately and severely disabled. The principal aim of special education is to help the disabled persons establish themselves in society through different special education programs depending on their degree of disability. The policy describes the special education strategy as: conducting national surveys on the prevalence of disability in accordance with types and degree of disability; improving the quality of existing special and integrated educational institutions and increasing the number of special and integrated schools for different types of disabled children; initiating an integrated education system in district and sub-district level primary schools; to establish teachers training colleges/institutions for teachers of special schools; to include disability issues in mainstream

teachers training curriculum; provision to be made for ensuring free supply of education materials to disabled pupils; alternative curriculum to be followed for children unable to cope with the mainstream curriculum etc. The National Education Policy (2000) does not include any specific policy guideline or action plan to either address or facilitate inclusive education. Rather, the emphasis is on special and integrated education. The strategies mentioned in the policy for special education, remain on paper and have not been implemented yet.

PEDP-II

Following recommendations made in a study in 2002 carried out by CSID in association with Cambridge Education Consultants Limited, UK, (commissioned by the Department of Primary Education, Government of Bangladesh), The Ministry of Primary Education in its Primary Education Development Project (PEDP) –II included a component of inclusive education for children with disabilities from 2004. However, it has not been implemented yet.

National Education Policy 2010

According to national education policy 2010, at present, the drop-out rate till or before the completion of Class V is about 50% and of the rest, about 40% leave the school before completing Class X. It is extremely urgent to bring down this rate of drop-out. So, necessary measures will be implemented so that all students are enabled to complete Class VIII and it will be ensured by 2018.

Physically Challenged Students

21. The facilities of the lavatories and the scope of smooth movement will be adequately designed and created with special attention in order to fulfill the special needs of the physically challenged learners.

22. Special and preferential attention will be given to their needs.

23. At least one trainer will be recruited in each of the PTIs to facilitate the special teaching methods and needs of various types of challenged learners.

Teaching Methods

26. An interactive teaching method will be pursued to develop the creative faculties and skills of the children and help them do the exercises through individual or group-work. Research initiatives to find out the appropriate methods for innovation of effective teaching, evaluation and implementation will be encouraged and supported.

Student Assessment

27. In Classes I & II, there will be continuous assessments, while from Class III onwards, quarterly, half-yearly & yearly examination systems will be in place. On the completion of Class V, a terminal examination with identical set of questions will take place at Upazilla/Pourashava/ Thana levels (of big cities). On the completion of Class VIII, a public examination will take place to be initially known as Junior School Certificate Examination. The Education Boards concerned with examination will conduct this public examination.

18. Special Education, Health & Physical Education, Scout, Girls' Guide and Bratachari

A. Education for Challenged Learners: Special Education

Aims and Objectives

• Steps will be taken to include the handicapped in the mainstream education.

• Special education will be provided to the acutely handicapped children who cannot fulfill the demands of daily life due to their physical or mental disabilities. These children are incapable of studying in the usual schooling system. Other than special education, they will be brought under efficient remedial system, special care and nursing.

Strategies

1. Survey will be conducted to find out the exact number, type of challenges and to categorize the handicapped population as per the degree of their disabilities.

8. Challenged children, unable to cope up with one or more than one subject, will be allowed to follow a flexible curriculum.

Based on the above stated literature review, the researcher as well as the readers could obtain a holistic concept on the legalize issues related to the need for developing an up - to - date assessment scale for Bangladesh.

CHAPTER THREE

Method

Process of Standardization

The present research was conducted in four steps as a part of standardization process of Stanford Binet Intelligence Scale. First step was the strength or weaknesses of items were identified through item analysis. Secondly, the norm was calculated and developed, thirdly reliability and afterward validity was tested. For different estimation, different samples were considered. For the calculation of norm, the study has considered students from six divisional metropolitan cities (Barisal, Chittagong, Dhaka, Khulna, Rajshahi and Sylhet) to represent Bangladesh. Instrument as SB5 for standardization (discussed in this chapter and briefly in chapter two) and standardized procedure for test administration (also outlined in this chapter) were followed for all the participants in different steps.

Research Design

The research was designed to fulfill the criterion for standardizing the cognitive ability test; the study was designed in the following four steps involved in standardization process.

Figure 4. Research design for the standardization of SB-5 in Bangladesh.

Item Analysis	• Cross sectional quantitative study • Expert opinion
Norm Development	• Cross sectional quanititative study
Reliability	• Repeated cross sectional quanititative study
Validity	• Cross sectional quanititative study

The following table summarizes the different process involved in standardizing the SB-5 on 11 age groups (06 – 16 years).

Table 3

Process in Standardization of SB5 for Bangladesh

Steps	Process	Participants	Area of Study	Total Participants	Time Period of Data Collection	Instrument Used
Step1	Item Analysis	30 from each age (11 age groups)	Dhaka division	330 (30x11)	June-Dec, 2005	Translated original SB5
Step2	Norm Development	50 from each age (50x11=550)	Six division	3300(550x6)	February, 06, August 07	Adapted items SB5 after item analysis
Step3	Reliability	30 from each age groups, twice (30x11) twice	Dhaka division	330 in 1st & 2nd administration	Jan - June, 2008	Same adapted test instrument was used twice with 7 days gap on same student
Step4	Validity	Expert opinion	Professionals	6	December, 2004	Review of Test Instrument
		20 special and 20 normal students from each age group	Dhaka division	440 (20x11=220) & (20x11=220)	July-Dec, 2008	Same adapted test instrument was used
		30 students form 3 age groups for both SB5 and WISC-R	Dhaka division (7, 11, and 14 years)	180 (30x3=90) & (30x3=90)	Jan - June, 2009	Both SB5-BD and WISC-R were used

Participants

As portrays above (Table 3), standardization is consisting of four steps, in different steps, this study considered different participants. The description of the selected participants and analysis for different steps is given below separately.

Steps of Standardization of SB5 for Bangladesh

Step I: Participants for Item Analysis of the Scale.

As the first step for standardization of an intelligence scale, each item of SB5 was analyzed through item analysis. To accomplish this objective, 30 students from each age were taken into account from 11 age group (6 to 16 years) totals of 330 students of whom 165 were male and 165 were female respectively. The participants for the item analysis were taken from Dhaka division only. For item analysis the following procedure is described.

Translation of Items.

The standardization process involves the adaptation of items based on the respective language and culture. In this context, the test instructions and various written activities of the original SB5 were translated into Bangla for participants understanding. The translation was done by graduate students of Institute of Education and Research (IER) twice under the supervision of experts (Appendix 4) for consistency. Then the final version was examined by six specialists [Professors of Department of Psychology and IER, University of Dhaka] to have consent on the quality and standard of translation. Three enumerators who were graduate students from IER, University of Dhaka were trained to translate the items for seven days. The translated items were pre-tested and the actual

administration of the test began with the direct supervision of the researcher. Finally, the translated version was retranslated into English by two Masters Students of IER separately in order to verify the reliability of translation. Considering the rapport building and the test administration, it took around one and half hour to conduct the test. Data collection for item analysis was conducted during June- December, 2005, where the researcher had to be considerate for the school examination schedule and vacation of the students.

Analysis Procedure of Items.

Both item Difficulty Index (P) and Discrimination Index (D) were calculated to determine the strength and weakness of each item. The Item Difficulty (P) was computed by calculating the total number of individuals who passed each item in the two different groups (i.e., upper and lower) divided by the total number of individuals who attempted the item from both groups for each item (Ahmann & Glock, 1981). Where Difficulty Index (P) was constructed following as:

Equation 1: **Difficulty Index (P)**

$$P = \frac{U_p + L_p}{U + L}$$

Item Discrimination or Discrimination Index (D) refers to the degree to which an item differentiates correctly among test takers in the behavior that the test is designed to measure. The purpose of this index is to provide a measure that would separate an item considering their correct or incorrect responses depending on the answers from upper and lower groups which explores discrimination of group differences. (Anastasi, & Urbina, 1997). Where Discrimination Index (D) was constructed following as:

Equation 2: **Discrimination Index (D)**

$$D = \frac{U_p - L_P}{U}$$

U_p=Number of students in upper group who answered the test item correctly
L_p=Number of students in lower group who answered the test item correctly
U= Total number of students in upper group
L = Total number of students in lower group

In this study, the item difficulty and the discrimination index were calculated within 165 children in total age group from the 330 students of the total sample. Following the total raw scores obtained by the students, scores were divided into two groups as upper group (first 25% students) and lower group (last 25% students). The students were ranked considering their total raw scores obtained through the test. According to the testing procedure, the students responded to items based on their age appropriate basal ability level. The item analysis included a total of 330 students, considering 30 students from each age group (6 to 16 years). These analyses were done into two ways. In the first stage, in calculating the difficulty index and discrimination index, participants from all age groups (6 to 16) were considered. In the second stage, age specific difficulty index and discrimination were calculated. For a binary response, score one was assigned for a correct answer and zero for a false answer. But for a multiple response where items were marked as zero, one and two; while two was considered as success (regarded as one) and rest areas failure (regarded as zero). However, the one from multiple answer contributed in calculating the final score, based on which score of the groups (upper and lower) were estimated.

Besides, Cronbach's alpha, the internal consistency or reliability of the test, is an overall item correlation. Cronbach's alpha reliability coefficient normally ranges between 0 and 1. However, there is actually no lower limit to the coefficient. When the Cronbach's alpha coefficient is closer to 1.0, the internal consistency of the items are greater in the scale. The Cronbach's alpha was estimated to find the internal consistency among the items. (Cortina, 1993). Where, Cronbach's basic equation for alpha was constructed as:

Equation 3: **Cronbach's alpha**

$$\alpha = \frac{n}{n-1}\left(1-\frac{\Sigma Vi}{Vtest}\right)$$

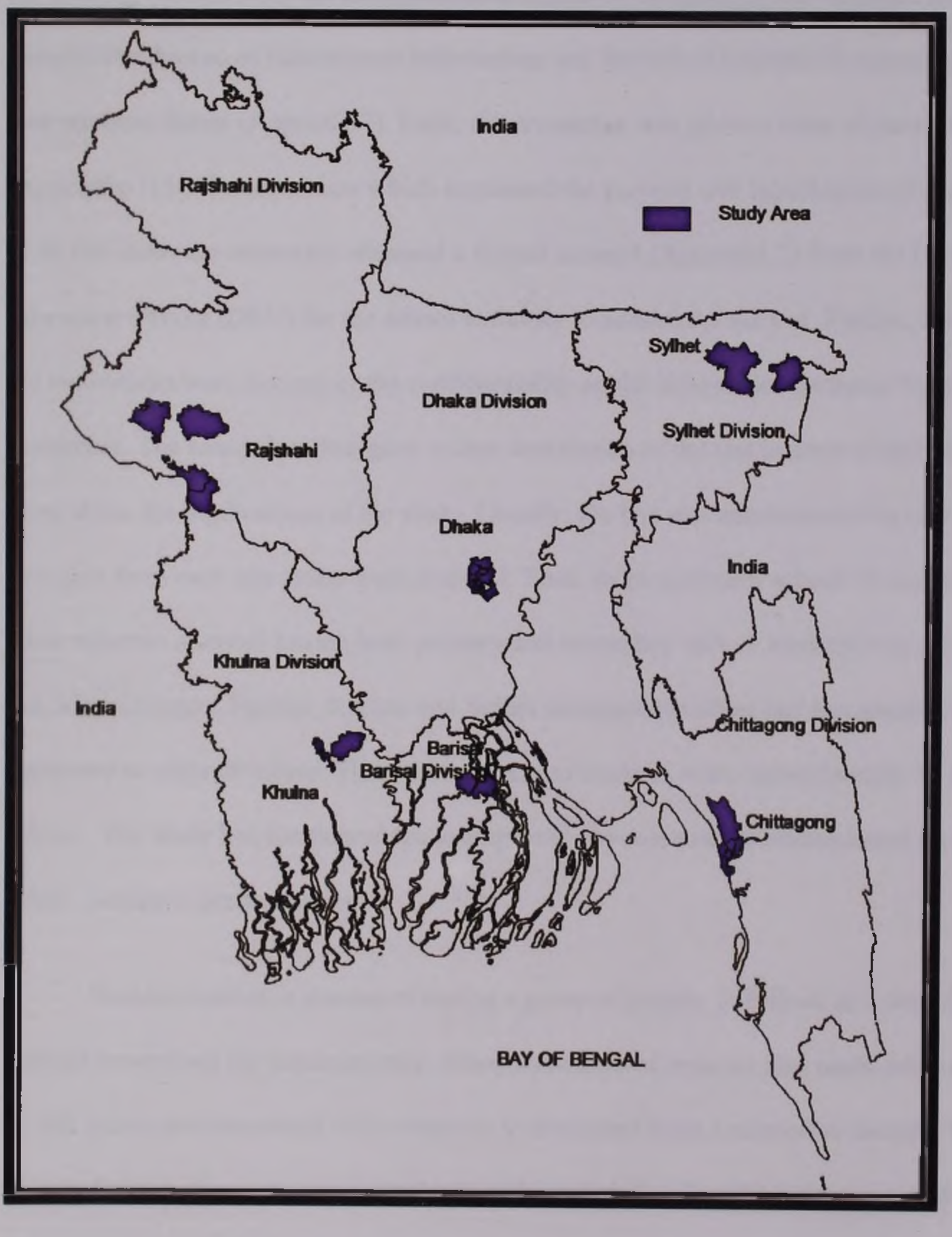
n = number of questions
Vi = Variance of scores on each question
V test = Total variance of overall scores (not %'s) on the entire test

Step II: Sampling for Construction of Norm.

The second step of this standardization process is the development of norms. A two-stage stratified sampling was followed in selecting sampling for construction of norm. In this regard, the whole of Bangladesh was treated as universe that contains six divisions. As Rangpur was separated as a division on January 25, 2010, hence this study did not consider this division as a sampling area (Appendix 5). Each division was considered as a stratum whereas the researcher has selected purposively the urban strata for her study. To represent the norm for Bangladesh, the study followed again stratified random sampling technique such as age, sex, geographic (divisional urban) region, socio economic level (only education of parents, each of which were defined in chapter two) which were considered as the stratified

variables for this study. Thus, the researcher considered 50 students from each age in 11 age group (6-16 years) from six divisional metropolitan cities (Barisal, Chittagong, Dhaka, Khulna, Rajshahi, Sylhet) total of 3300 (50*11*6) students (Appendix 6). After item analysis the adapted items were utilized for the study. The data collection was conducted during February 06- August 07. For the norm calculation, the researcher herself collected all the data of 3300 students from the six divisions, where she had been accompanied and supported by an assistant. This study included more than 150 schools (Appendix 7) from the divisional metropolitan cities of Bangladesh which were purposively and randomly selected (Figure 5: Map 1).

Figure 5. Study Area (Bangladesh Map).



General Procedure for Field Study for Norm.

Prior to the task of testing, the following procedure was taken into account. A list of the existing schools from six metropolitan divisions of Bangladesh was obtained from the Bangladesh Bureau of Educational Information and Statistics (BANBEIS) representing the metropolitan thanas (Appendix7). Later, the researcher was given a letter of permission (Appendix 1) by the supervisor which explained the purpose and significance of the research. With this letter the researcher obtained a formal consent (Appendix 2) from the District Education Officer (DEO) for the school authority to administer the test. Further, the head of the institutions were assured of the confidentiality on the information gathered by the researcher. The researcher then gave a clear description of the test in general and shared with them about the implications of the study. Usually, the test was administered on two boys and two girls from each age group from a school. Thus, from a primary school 20 students were taken whereas a school having both primary and secondary unit 44 students was selected for test administration. Further, Khulna and Sylhet metropolitan cities had few number of schools compared to other divisions. Thus, the numbers of students were approximately 50 to 55 per school. The study has considered the average and above average students based on their school academic performances.

Standardization, a process of testing a group of people, is defined as a test with clearly defined procedures for administration. Many standardized tests are also norm referenced; that is, test scores are interpreted with reference to the scores from a normative sample. With standardization, the norm group must reflect the population for which the test was designed. The group's performance is the basis for the tests norms. Standardized testing involves using testing instruments that are administered and scored in a pre-established standard or consistent manner. For the present study age norm was computed from the raw scores

obtained by the subjects after administering the nonverbal and verbal subtests (five subtests from each domain) of the Stanford-Binet Intelligence Scale (Fifth Edition).

In order to construct the age specific norm from the raw scores, the following steps were applied.

- Conversion of Raw Scores to Scaled Age Score (SAS)
- Conversion of SAS to Non-verbal and Verbal SAS
- Conversion of Nonverbal and Verbal IQ to Full Scale IQ (FSIQ)

Conversion of Raw Scores to Scaled Age Score (SAS).

In the first step, frequency distributions were made for each age group and each sub test separately. Later, mean and standard deviation (SD) was calculated with the raw scores for the ten subtests. Then small z was calculated. Usually, the mean and SD of each subtest were computed through-

Equation 4: **Small z Score**

$$z = \frac{x - \bar{x}}{SD}$$

Where,

x = Raw score
$\bar{x}$ = Mean of the raw scores
SD= Standard Deviation derived from the raw scores.

And then the big Z was calculated with each individual small z scores. In this computation, the weighted mean and the standard deviation were measured.

Equation 5: **Big Z score**

$$Z = Sz + M$$

Where,

S= Weighted Standard Deviation
z= Small z scores
M= Weighted mean

Z scores are a type of standard score. The z score is useful when attempting to compare items from distributions with different means and standard deviations. The z score for a test score indicates how far and in what direction that test score is from its distribution's mean, expressed in units of its distribution's standard deviation. The z scores will have a mean of zero and a SD of one. To estimate the z score from the raw scores, the researcher need to calculate the weighted means and SD from the raw scores. Further, with the obtained mean and standard deviation from the raw scores, weighted mean and standard deviation was again estimated for both the non verbal and verbal subtests of the scale.

Conversion of SAS to Non-verbal and Verbal domain SAS.

For the calculation of Scaled Age Score (SAS), a frequency distribution was again calculated with the sum of subtest SAS. Further, again small z and big Z were calculated for each verbal and non verbal SAS. The weighted mean and standard deviation were applied to calculate the big Z.

Conversion of Non-verbal and Verbal domain to Full Scale (FSIQ).

The sum of the two domains SAS were again calculated to find out the mean and the standard deviation as before the small z and big Z were calculated. The weighted mean and Standard Deviation (SD) were again applied for the calculation of Full Scale IQ.

Step III: Participants for Determining the Reliability.

Test retest reliability was computed to determine the reliability of SB5. The most common method for finding out reliability of test score is by repeating the identical test on a second occasion. The reliability coefficient in this case simply the correlation between the scores obtained by the same students on two administration of the test based on the norm developed from this study. Test-retest reliability shows the extent to which score on a test can be generalized over different occasions. The higher the reliability, the less susceptible the scores are to the random daily changes in the conditions of the subject or of the testing environment (Anastasi, 1992). Then Pearson's product moment formula (correlation) was quantified to determine test retest reliability.

Equation 6: **Correlation Coefficient**

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$

Where,

r = Reliability correlation coefficient
X= IQ score in first administration
Y= IQ score obtained after the second administration.
n= number of children

In step III, the ten subtests of the Stanford-Binet Intelligence Test were administered twice with an interval of one week on 30 students from each age and from 11 age group (6 to 16 years) to determine the test retest stability of the adapted version. Hence, a total of 330 participants were taken twice from Dhaka division. The data collection was done during Jan - June, 2008. The researcher herself administered the test along with the assistant. Since the

test was administered twice, there was a difference in testing time as on average one hour for first administration while about 45 minutes for the second administration.

Step IV: Participants for Determining the Validity.

In this study the validity for SB5 was established by the determination of content related, criterion related evidence. It is vital for a test to be valid in order for the results to be accurately applied and interpreted. Validity of test scores depends on the proper administration of the test by an experienced examiner and validity is not determined by a single statistic, but by a body of research that demonstrates the relationship between the test and the behavior it is intended to measure (Cronbach, 1970).

Content Validity.

Content Validity is based on the extent to which a measurement reflects the specific intended domain of content (Carmines & Zeller, 1991). Content related validity was done by the consultation with professionals.

Consultation with Experts.

Intelligence testing in modern society has changed in many respects from the traditional uses of testing in the past. One major change in modern perspectives of intelligence testing is that IQ is now viewed to be influenced by both heredity and by environmental effects. (Dickens & Flynn, 2006). Schooling, in particular, has large and beneficial influences on human intelligence (Ceci,1994). Many types of intelligence tests are used in schools in today's world. A number of intelligence tests are currently, being administered to children in the early grades. Thus, educators of modern society now a day

emphasize the need for appropriate testing tools in making long term ability based decisions for student's academic progress (Kolar, 2001).

Considering the above issues, the SB5 (Roid, 2003) was given to six experts for review and content validation of the items on December 2004. The professionals were from the Department of Psychology and Department of Special Education, Institute of Education and Research (IER), University of Dhaka (Appendix 4). Formal discussion was made on the items of SB5 and pattern of norm with different professionals of this area. Finally, the professionals came to a consensus and recommended the standardization of SB5 and test adaptation for use in urban Bangladesh to measure the cognitive and intellectual abilities of children.

Criterion Validity.

Criterion validity is used to demonstrate the accuracy of a measure or procedure by comparing it with another measure or procedure which has been demonstrated to be valid. (Carmines & Zeller, 1991). The following procedures were used in computing the criterion validity.

- Contrasted Group Validity
- Correlation with the Wechsler Scale and SB5-BD

Contrasted group of validity.

A total of 440 students participated from Dhaka divisional schools for the study. Among them 20 from each age were considered with a total of 220 students from 11 age group (6-16 years) and similarly 220 students with special needs. The adapted version of the

scale was administered to determine the validity on two groups of students and then both the IQ scores were compared to validate the scale. On average, the time duration for test administration was around 45 minutes and 60 minutes for special needs and normal students respectively. For this purpose, the study was worked out during July-Dec, 2008.

When referring to validation by contrasted groups, Anastasi states," in the validation of an intelligence test, the scores obtained by intellectually disabled children may be compared with those obtained by school children of the same age" (Anastasi, 1997).

The following'**t**' test was computed by comparing the two mean IQ scores of normal and special needs student. The'**t**' test formula was-

Equation 7: '**t**' **test**

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{SS_1 + SS_2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Where,

$\bar{X}_1$=mean IQ score of normal student
$\bar{X}_2$=mean IQ score of special needs student
n_1= number of normal students
n_2= number of special needs
SS_1= sum of squares of deviations from $\bar{X}_1$
SS_2= sum of squares of deviations from $\bar{X}_2$

Correlation with the Wechsler Scale.

In order to validate the SB5, the scores of 90 students obtained by SB5 from three age groups (age 7,11 and 14 years) were correlated with the scores of same 90 students who were administered the Wechsler Intelligence Scale for Children-Revised (WISC-R). On average, the time duration for test administration was around 60 and 75 minutes of SB5 and WISC-R

respectively. And correlation between the IQ scores obtained by this study and WISC-R (Huq, 1980) (description of WISC-R were presented at the end part of this chapter) were compared to see criterion related validity. For this purpose, the study was worked out during Jan - June, 2009.

Instruments Used

This study used two types of tests, firstly, the SB5 scale for standardization and secondly, WISC-R to justify the criterion validation of SB5.

Testing Standards for Selection of a Standardized Test

The history of standardized assessment in the United States is longer than that of alternative approaches; many standardized tests are readily available. The Standards for Educational and Psychological Testing is a set of testing standards developed jointly by the American Educational Research Association (AERA), American Psychological Association (APA), and the National Council on Measurement in Education (NCME). According to Testing Standards of AERA, APA & NCME (1999), the following criteria can be considered in selecting a test that is appropriate for the child being assessed. The criteria for testing selection are : a detailed and comprehensive written procedures in test manual , development of a large and diverse normative sample, adequate skills on sampling, collection of current normative data, strong reliability and validity, a clear instructions for the administration and scoring of the test, appropriate stimulus items, revised recently, familiar to test administration and yield useful diagnostic information and help to design treatment goals.

So, Stanford Binet Intelligence Scale Fifth Edition (SB5) fulfills all the criteria of standardized test that is used worldwide and selected for assessing intellectual ability of an individual.

The Description of Stanford-Binet Intelligence Scale Fifth Edition, 2003.

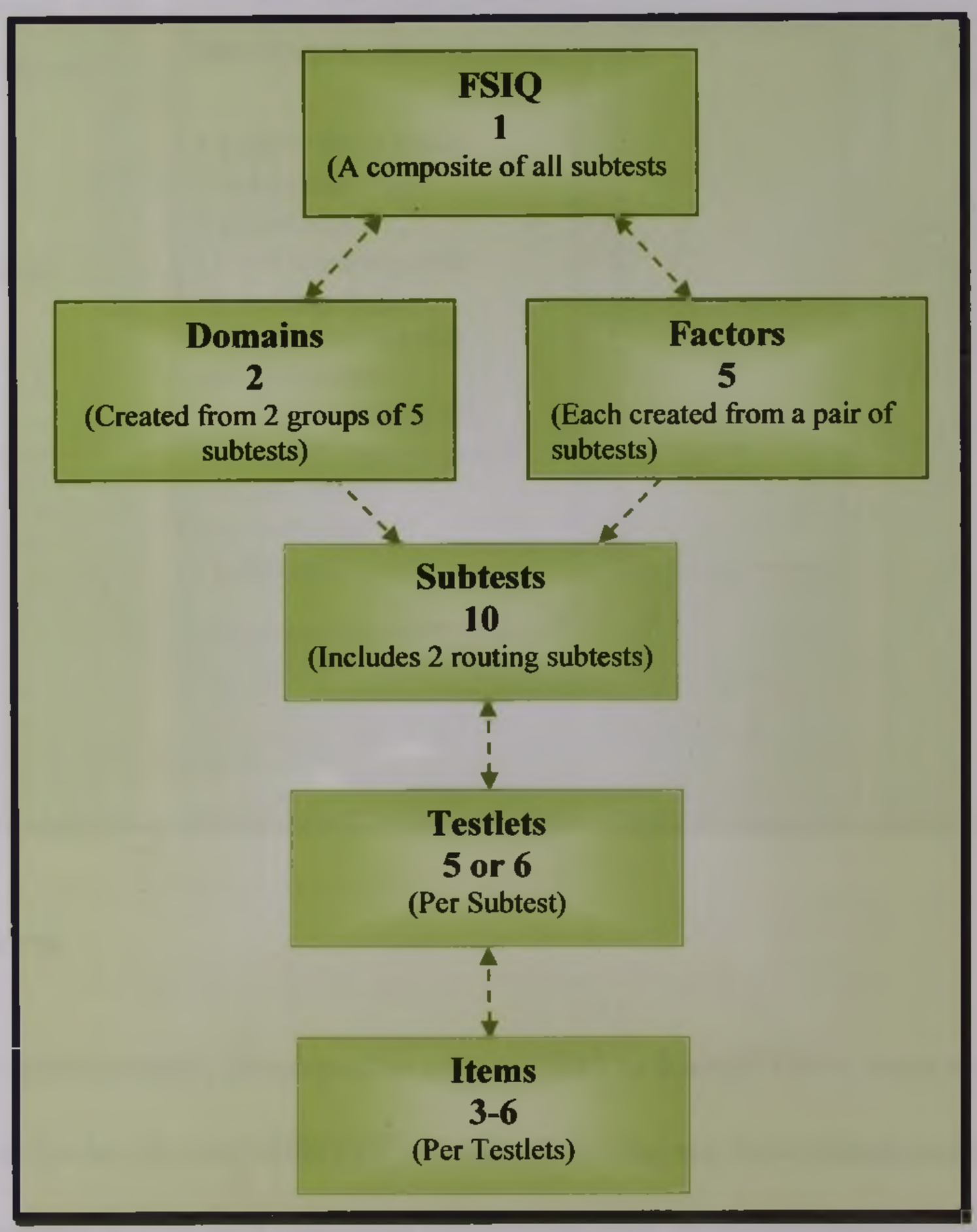
The Stanford-Binet Intelligence Scale, Fifth Edition (SB5) is a current assessment with a rich tradition. This new editions of nationally standardized tests provide modern wording, illustrations, enhanced measurement procedures, updated theory and research, and new standardizations, enhancing the validity of test interpretations. After a 7-year revision project, the Stanford-Binet Intelligence Scale, Fifth Edition (SB5) (Roid, 2003b) was published with enhanced features, norms and procedures. It blends many of the important features of earlier editions (all editions described and reviewed in chapter two) with significant improvements in psychometric design. It incorporates the use of two routing subtests in the point-scale format of the 1986 edition with the functional-level design of the 1916 to 1960 editions for the remaining subtests. By adapting the test to the functional level of the examinee, the routing procedure of the SB5 increases the precision of measurement by tailoring the difficulty of the items to his or her level of cognitive functioning. The use of a hierarchical model of intelligence with a global g factor and several broad factors at a second level is repeated in this edition. (Roid, 2003)

A Hierarchy of Components from Items to Full Scale IQ.

The items of SB5 at the basic level consist of individual tasks or problems that are scored for pass or fail. Figure 6 (below) shows the SB5 incorporates items into more general components. Groups of three to six items combine to form testlets. Items within a testlets are at a common range of difficulty. Groups of five to six testlets, each at increasing levels of

difficulty, combine to form a subtest. Subtests are combined into either one of the two domains or one of the five factor indexes. At the most general level, either the two domains or the five factor indexes combine to form the Full Scale IQ (FSIQ). The Full Scale IQ derives from the administration of ten subtests and is considered the standard measure of global intellectual ability. The ten subtests used in SB5 are summarized in the following figure according to the five factors and two domains.

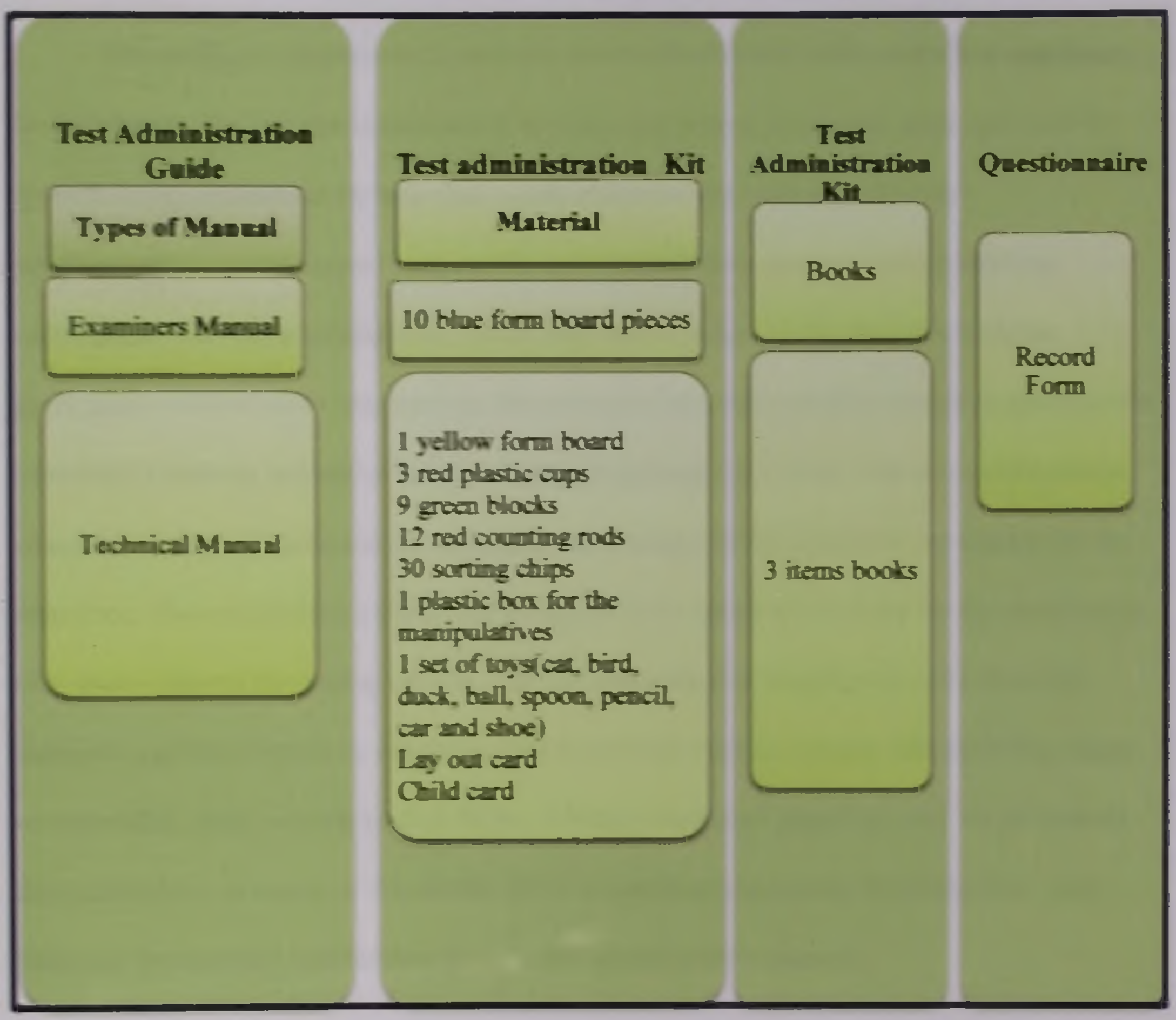
Figure 6. Hierarchy of Components in the SB5.



Source: Roid (2003), p.25.

Figure 7. Materials in the test kit.

The complete SB5 test kit contains the materials listed below.



Record Form.

Structured questionnaire, developed in original SB5 as Record Form, were applied to judge and calculate the intellectual ability of individuals for the standardization process. The supervisor procured the test kit with copyright permission from riverside publishers along with the software CD to calculate the IQ following the norm developed by SB5 original (Roid, 2003).

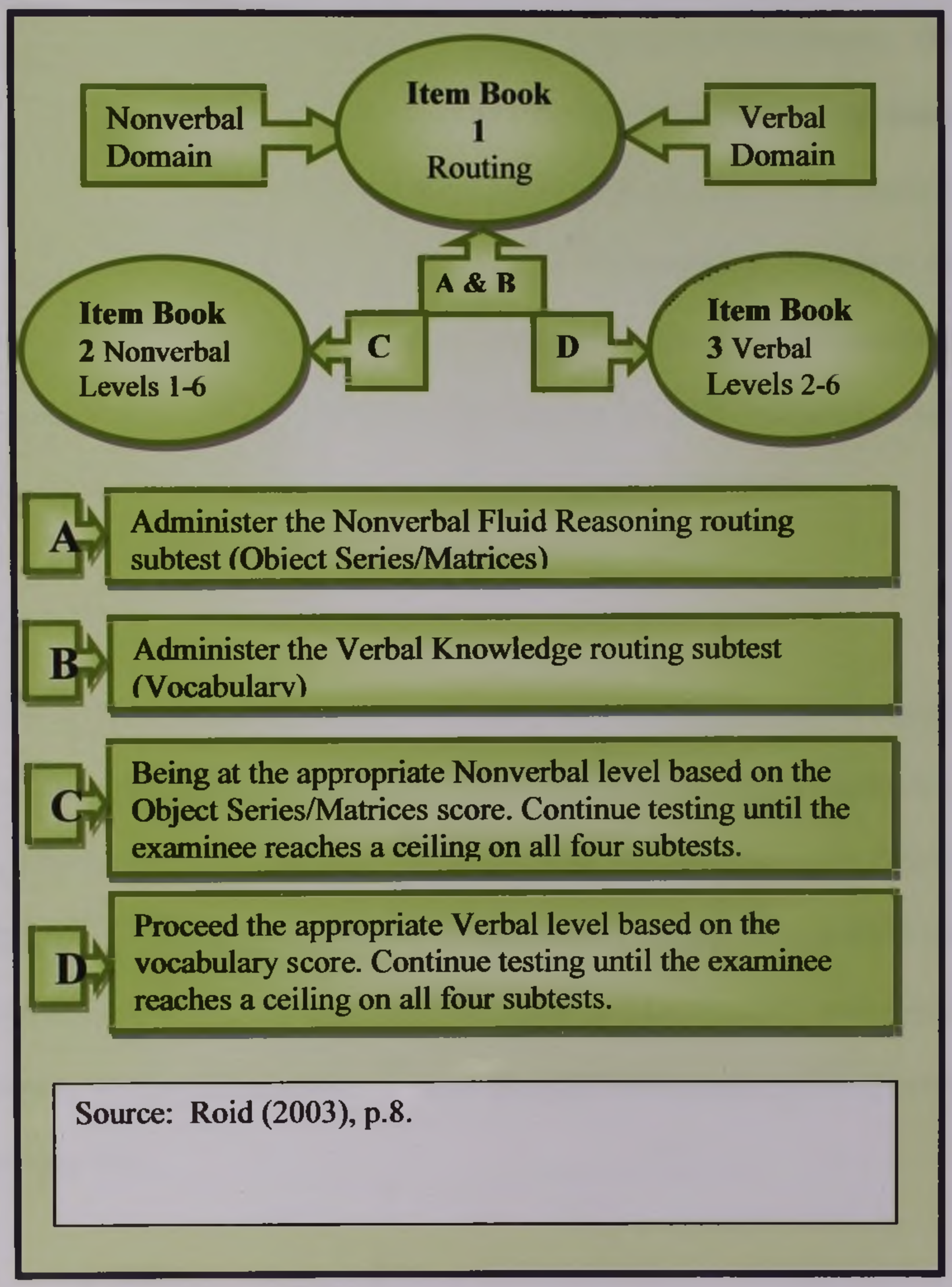
Procedure

Standardized Testing Procedures for SB5.

Standardized test assesses a student's functional abilities under controlled conditions. In this context, the test was administered in a separate setting along with quiet and well-lit room to avoid extraneous variable that would affect the test scores. Before test administration, a good rapport between the test administrator and the participants was established. The test administrator (researcher) was sensitive to the pace at which the participants worked most comfortably. She presented the tasks rapidly enough to maintain the examinee's interest, but not so quickly that the examinee felt rushed. She also established a relaxed and pleasant environment and made the testing session a positive experience for the examinee. The record form of SB5 also comprised of a series of checklist for the observation of a student during the testing session. The test administrator would note in the form any unusual examinee responses, reactions, or distractions, such as extreme distractibility, anger or opposition, poor communication skills, or highly emotional responses and would include this information as report of the results. To be scientifically accurate, the researcher also followed the standard instructions given in the administrative manual.

The Standard Order of Administration for the ten subtests comprised of a systematic layout as proposed by the original scale was also followed for the testing sessions (Figure 8).

Figure 8. Standard administration order for the Stanford-Binet Intelligence Scales, Fifth Edition.



Routing Subtests

Two special subtests—the routing subtests—were administered at the beginning of the SB5. The routing subtests identify an individual's developmental starting point for the entire remaining subtest.

The Nonverbal Routing or Fluid Reasoning subtest (Object Series/Matrices) provides an indicator of an individual's nonverbal ability and serves as the basis for determining his/her starting point for the remaining four subtests in the nonverbal domain. The other routing subtest, Verbal Knowledge (Vocabulary), provides an indicator of an individual's verbal ability and is the basis for determining his/her starting point for the remaining verbal subtests. Both of the routing subtests are included in SB5 Item Book 1 and Item Book 2 contains the remaining nonverbal subtests, and Items Book 3 contains the remaining verbal subtest. Figure 8 shows this organization and the proper order for standard test administration for an explanation of an alternative nonverbal administration of the SB5.

Levels.

Items for all subtests except the two routing subtests are grouped into testlets. These testlets are then arranged into levels of difficulty, with six levels for the Nonverbal domain and five levels for the Verbal domain. In both domains, Level 6 is the most difficult, but the Verbal domain only contains five levels, labeled Levels 2 through 6. Nonverbal Level 1 consists of two testlets at the lowest level of difficulty and has no direct counterpart in the Verbal domain. Within each domain, Levels 2 through 6 each consist of four testlets, one for each remaining factor.

Activities.

Many of the SB5 subtests contain more than one type of item. This is necessary because of the wide range of ages and abilities that each subtest spans. An activity that works well to assess a particular factor for young children may not be the most appropriate way to assess that factor for adolescents or adults. For example, the Nonverbal Visual-Spatial

Processing subtest uses simple Form Board activities for the initial tasks and Form Patterns activities for all tasks at subsequent levels.

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

The Wechsler Intelligence Scale for Children -Revised (1974) was used to validate the SB5 by computing the correlation among the three (Verbal, Nonverbal and Full Scale) IQ scores. The scale was translated and adapted in Bangla in 1980 by Sharmin Huq. The age ranges were from 6 -15 years. The WISC-R comprised of the verbal and the performance scale. Each scale consisted of five subtests. The ten subtests are shown through the following Table 4.

Table 4

Showing the ten subtests of WISC-R (Bangla version)

Verbal Subtests	*Performance Subtests*
Information	Picture Completion
Similarities	Picture Arrangement
Arithmetic	Block Design
Vocabulary	Object Assembly
Comprehension	Coding

Ethical Consideration

A prerequisite of any standardized intelligence test is to follow the standard procedure as stated in the examiners and technical manual. Similarly, given the long history of Stanford Binet Intelligence Scale and the importance of accurate assessment of intellectual abilities, the researcher of this study followed the same instructions as proposed by the author (Roid, 2003) of original SB5. Considering the professional and ethical issues related to the overall assessment of IQ, researchers expertise, training, data collection process and its confidentialities, technical qualifications of scoring procedures, data analysis were regarded from the very beginning till the writing of the report. In this context, scores obtained from tests of intellectual ability such as the SB5 had to be as precise as possible, given that they were used for life changing decisions of treatment, placement or classification. However, the concept that all test scores had some degree of measurement error is critically important to the ethical use of tests. A shorter retesting interval would allow the SB5 to be highly useful in the assessment of treatment interventions in clinical and neuropsychological settings as well as in re-evaluations for special education. The stability of the SB5 is even more impressive in light of the relatively shorter test-retest interval on the SB5 (5 to 8 days) compared to that on the Wechsler scales (23 to 35 days on average). In addition, in this research, the researcher would also concentrate on the basic ethical norms required by the American Psychological Association (APA, 1992, 2000) during research study. Furthermore, naturally, researcher also followed an ethical obligation to prevent physical and mental destruction to her subjects. Researchers also would allow her participants to withdraw from the study at any time if they wish to stop participating. Finally, researcher pursued the strategy as an obligation to protect the ambiguity of their participants understanding on overall test administration.

CHAPTER FOUR

Results

The results section describes the findings of different segments that involved in completion of standardization process. Eventually, this section is also organized into four segments such as item analysis, standardization of norm & IQ of SB5, reliability and validity of the test.

Item Analysis

When norm-referenced tests are standardized for instructional purposes, to assess the effects of educational programs, or for educational research purposes, it becomes very important to conduct item analyses. Item analysis, as a first step of standardization of intelligence scale, was carried out through SB5 test kit among the 330 students of 11 age levels (6-16 years) to scrutinize the strengths and weaknesses of the test items. Examination of each item was done in terms of (i) Difficulty Index (ii) Discrimination Index. Item analysis was computed on the scores obtained by the participants for the ten subtests of SB5. As discussed earlier, item difficulty and discrimination index were calculated within 165 of the total sample group. Following the total raw scores obtained by the students, scores were divided into two groups as upper group (first 25% students) and lower group (last 25% students).According to the testing procedure, the students responded to items based on their age appropriate basal ability level. Item analysis results have been analyzed in the Table 7 and Table 8 showing the re-arranged items of the ten subtests of Stanford-Binet Intelligence Scale (Fifth Edition), Moreover, modified and adapted items for both verbal and nonverbal subtests based on P and D value are also presented. Besides, pictorial presentation of those modified and adapted items are shown in Appendix 10.

Difficulty Index.

The level of difficulty of an item focuses on the proportion of students who correctly answer an item. The higher the correct response from both groups, the easier the item. On the other hand, as the item gradually becomes more difficult the proportion of answering an item correctly becomes lower (Ahmann and Glock, 1981). An item difficulty of 1.0 indicated that everyone answered correctly, while 0.0 means no one answered correctly. The item analysis considered 30 students from each age group (6 to 16 years). These analyses were done into two ways. In the first stage, in calculating the discrimination index and difficulty index participants from all age groups (6 to 16) were considered. In the second stage, age specific difficulty index and discrimination were calculated. In calculating difficulty index and discrimination index 165 students from all ages were considered, where, 82 were from lower score group (lowest quartile) and 82 from upper score group (highest quartile).

Table 5

Difficulty Index (P) for each items of SB5 in Bangladesh (All age)

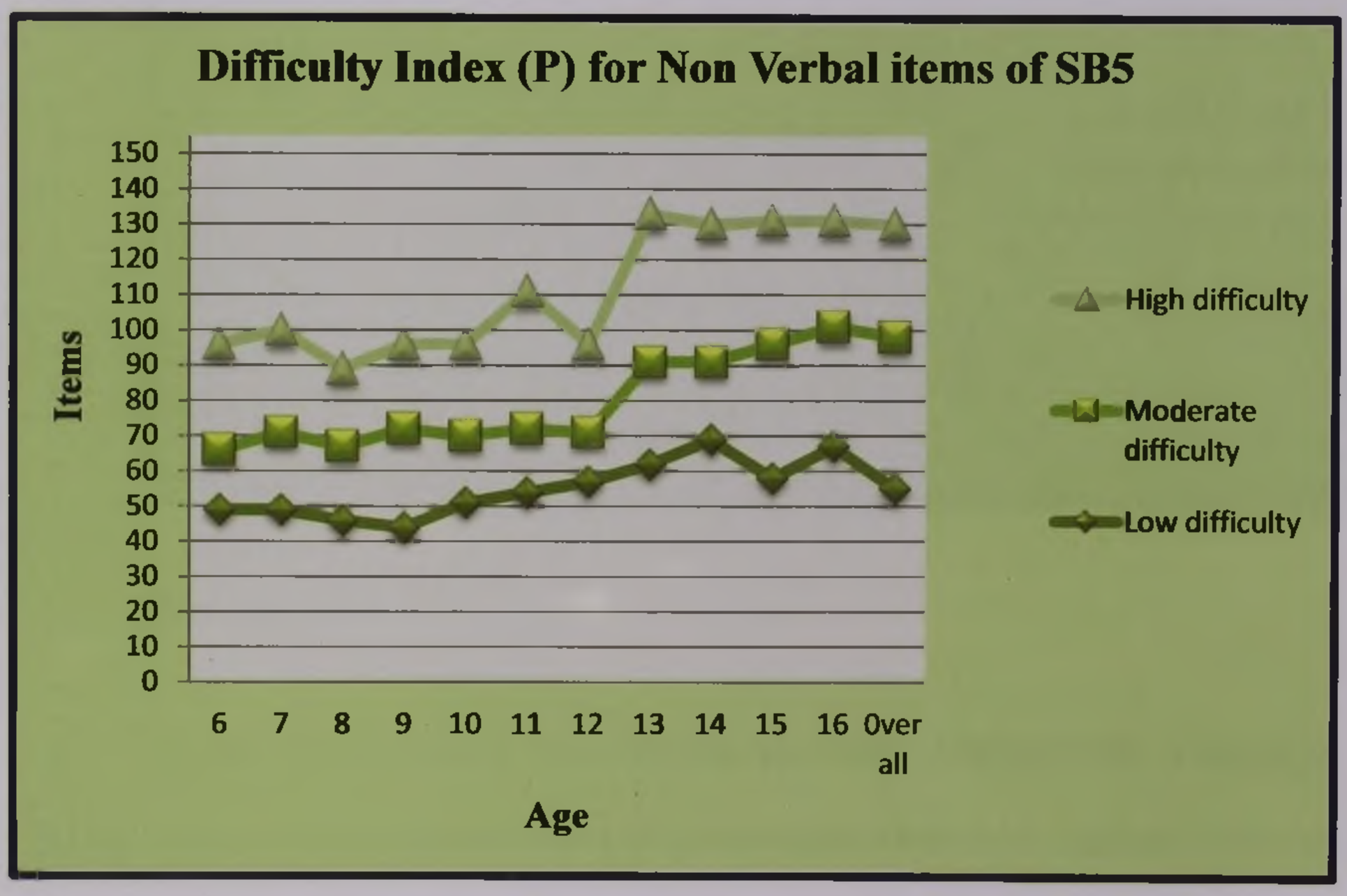
Difficulty Level	Difficulty Index (P) Value	Results from the Study	
		Non Verbal	Verbal
Low	Greater than and equals to 0.80	55 items	50 items
Moderate	Ranges from 0.31 to 0.79	43 items	54 items
High	Less than and equals to 0.30	32 items	15 items
	Did not answered	22 items	22 items
Total		152 items	141 items

The above Table 5 describes the number of items that were found difficult (low, moderate and high difficult) considering all age group. Findings reveal that from non verbal domain 55 items were found low difficult, 43 were moderate and 32 items were high difficult. From the verbal domain, 50, 54, 15 items were found low, moderate and high difficult respectively. It was also found that the difficult items were usually from upper level

of the test domain (e.g. level 5; level 6). Item-wise difficulty index value was presented in Appendix 9.

For age specific difficulty index, score and items of 30 students of each age were considered. Each age level was separated for the analysis. At each age level the study had 30 participants. Participants were ranked following their total raw score then grouped into higher and lower category. In this calculation, 15 students were from lower score group and 15 were from upper score group of the 30 students.

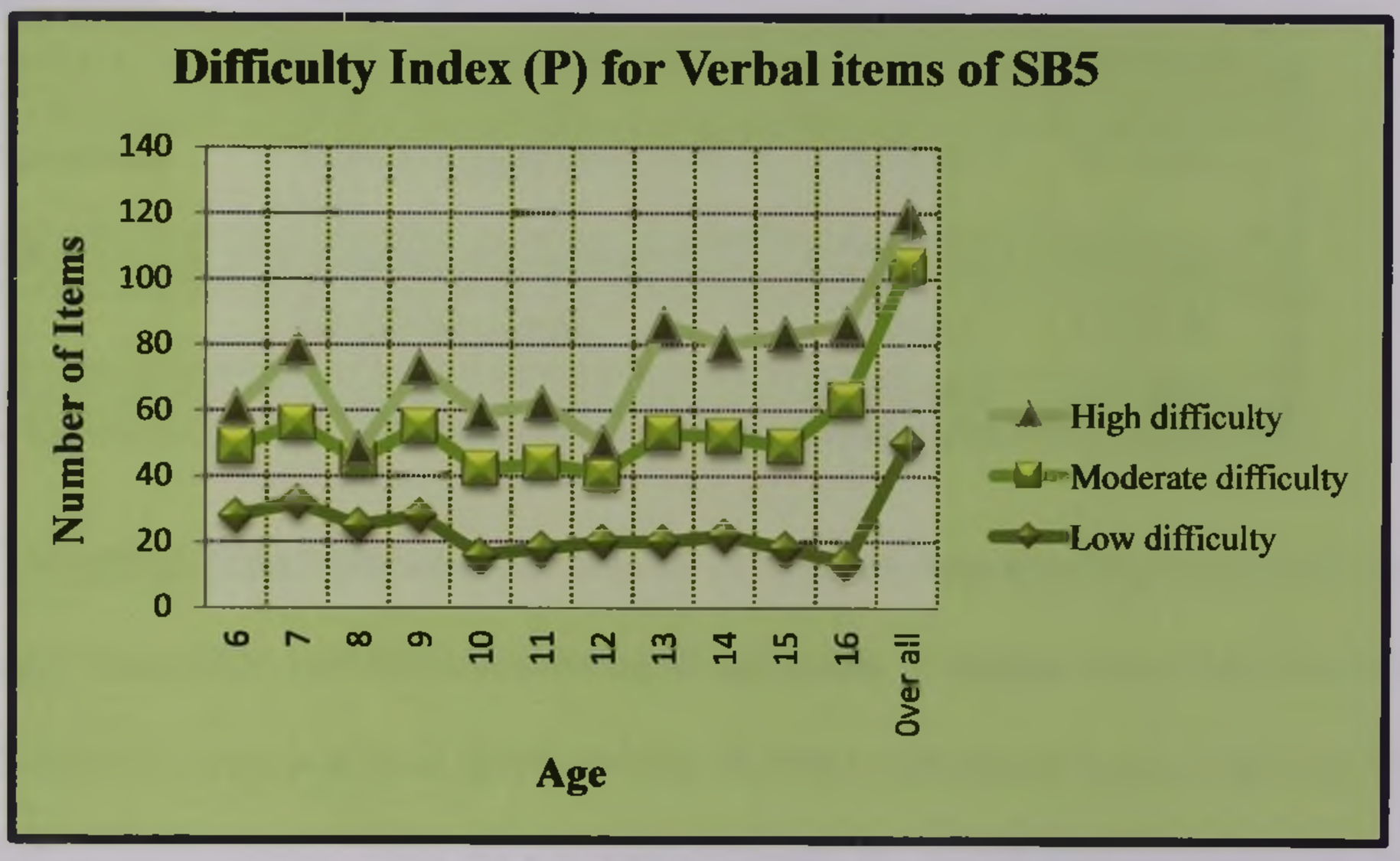
Figure 9. Difficulty Index (P) for Non Verbal items of SB5 in Bangladesh (Age specific).



The above Figure 9 shows that in non verbal domain as age increases, it reduces the number of unanswered items. Again, as age increases the proportion of low difficulty index increases except for age 15 years in non verbal domain. And the same trend was found in

verbal domain of the SB5 in Figure 10. However, higher portion of low difficult items were found in verbal domain. For age 15 years the low difficulty index did not increase in verbal domain. The students of Bangladesh show difficulty to gain the skill in verbal fluency. Along with this, the socio-cultural and education system might act as barrier to express their verbal views spontaneously.

Figure 10. Difficulty Index (P) for Verbal items of SB5 in Bangladesh (Age specific).



Discrimination Index (D).

The item discrimination index (D) can vary from -1.00 to +1.00. A negative discrimination index (between -1.00 and zero) results when more students in the lower group answered correctly than students in the higher group. A discrimination index of zero means equal numbers from higher and lower students answered correctly, so the item did not discriminate between groups. A positive index occurs when more students in the higher group answer correctly than the lower group (Jean-Marc, 2008). The following table depicts the items which have discrimination index and which do not have. As like as the difficulty index,

the discrimination index were analyzed into two sections namely all ages and age specific sections.

Table 6

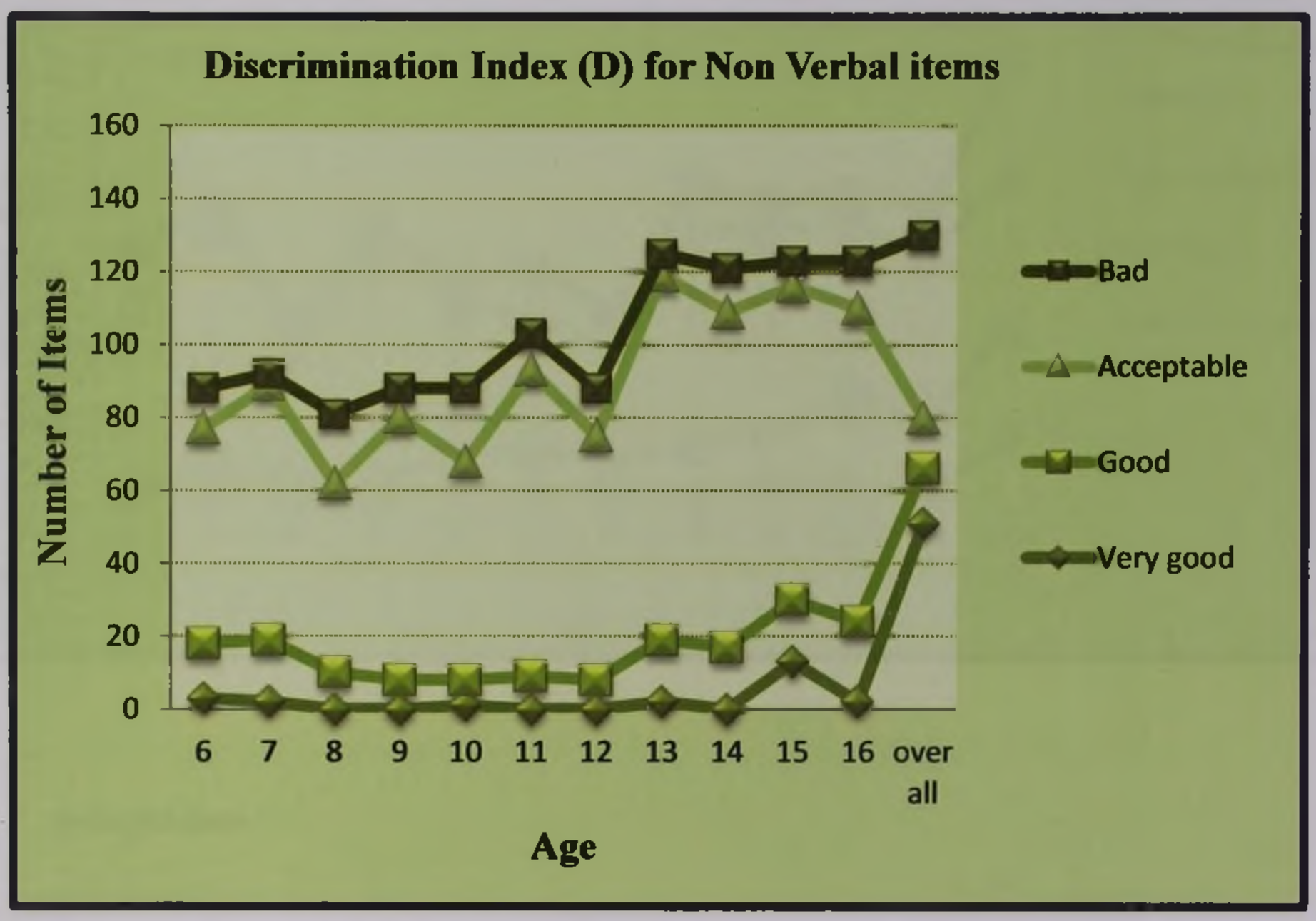
Discrimination Index (D) for each items of SB5 in Bangladesh (All age)

Discrimination Level	Discrimination Index (D)Value	Findings from the Study	
		Non Verbal	Verbal
Very good	Greater than or equals to 0.6	51 items	16 items
Good	When D ranges from 0.31 to 0.60	15 items	31 items
Acceptable	When D ranges from 0.01 to 0.30	14 items	41 items
Bad	Less than or equals to 0.20	50 items	31 items
	Did not answered	22 items	22 items
Total		152 items	141 items

The above Table 6 describes the number of items by their discrimination level (very good, good, acceptable, and bad) considering all age group. Findings reveal that from non verbal domain 51 items and from verbal domain 16 items were found bad considering the discrimination of the items. From the non verbal domain 15, 14, 50 items and from the verbal domain 31, 41, 31 items were found very good, good and acceptable respectively at the decimation level. It was also found that the discriminating items are usually from middle level of the test domain (e.g. level 3; level 4). Item-wise discrimination index values are presented in Appendix 9.

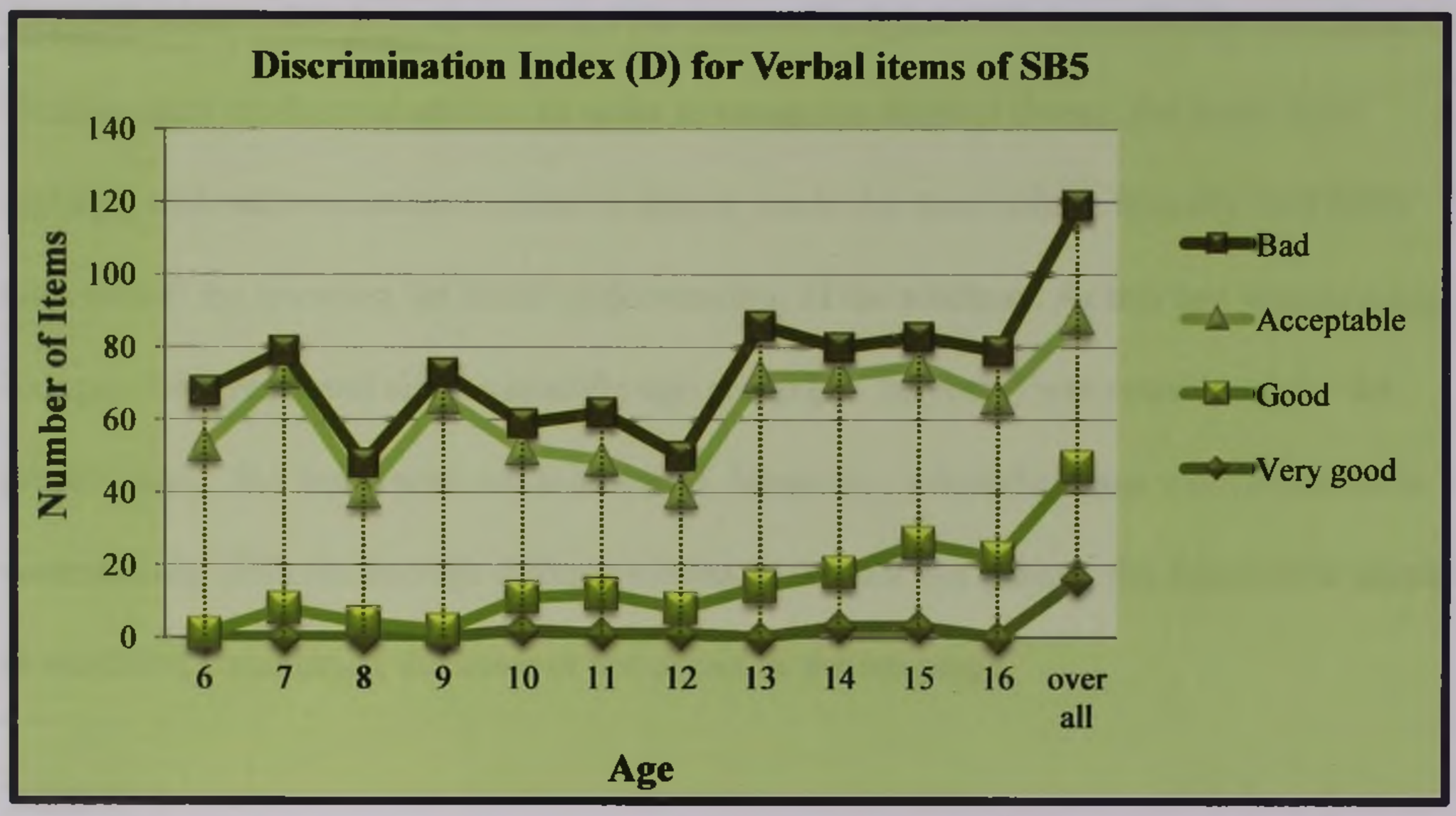
Similarly, for age specific discrimination index, score and items of 30 students of each age was calculated. The below figure 11 shows that in non verbal domain as age increases, it gradually reduces the number of unanswered items. Again, as age increases the proportion of acceptable items also increases.

Figure 11. Discrimination Index (D) for Non Verbal items of SB5 in Bangladesh (Age Specific).



And the similar trend was found in verbal domain of SB5 (Figure 12). A few numbers of items were found under good discrimination area in both non verbal and verbal domain. But during the all- age analysis (Table 6) a significant number of items were found to be good and very good discrimination showing that, the test has the quality and ability of item fairness.

Figure 12. Discrimination Index (D) for Verbal items of SB5 in Bangladesh (Age Specific)



Adaptation

As discussed in earlier chapter two, the test adaptation is a process by which a test (or assessment instrument) is transformed from a source language and/or culture into a target language and/or culture. The dynamic strength behind test adaptation is test validity (Geisinger, 1994cited in Matthews, 2003). Since we know that the purpose of any testing is to produce meaningful and interpretive assessment outcomes, then the aim of any test adaptation is the same; to provide a fair, equivalent, applicable and interpretable assessment instrument (Misra, Sahoo & Puhan, 1997 in cited Matthews, 2003). In accordance with this point of view, this research includes test adaptation for the completion of standardization of Stanford Binet Intelligence Scale Fifth Edition for use in Bangladesh. The standard guideline recommended by the International Test Commission (ITC) was followed for the process of adaptation (Guideline D1., ITC, 2001; Hambleton, 2005).

The adaptation and modification of the items were done based on the difficulty and discrimination index. It is obvious that the items of original SB5 thematically correlated to identify one's intellectual ability. In order to retain the original theme, the items were replaced with native content/symbol or object, made the item culture friendly, and often retranslated the question for better understanding of the students. As this test was developed for ages 2-85 years and since a specific age group (6 - 16 years) was considered for the present study, the items were not eliminated. Moreover, when the items were found to be continuously difficult through difficulty index in several age groups, the items were adapted or modified considering the color or the object or the language.

Table 7

Modified and adapted items based on P and D value (Non-Verbal Domain)

Item No.	Original Item	Replaced Item	Reason	Difficulty Index (P)	Discrimination Power (D)
FR_8	7 and 8	7 was placed at 8 and original 8 was replaced at 7	Original 7 was more difficult than original 8	0.73	-0.13
FR_27	26 and 27	26 was placed at 27 and original 27 was replaced at 26	Original 27 was more difficult than original 26	0.66	-0.20
FR_28	28 and 32	28 was placed at 32 and original 32was replaced at 28	Original 32 was found easier than 28	0.17	0.10
KN2_1	Feeds child	Theme unchanged	Replaced with local picture	1.0	0.64
KN3_2	Drinks with straw from glass	Bottle with straw	Culturally appropriate picture introduced	0.47	-0.60
KN3_3	Eat with spoon from bowl	Eat with spoon from a plate	Culturally appropriate picture introduced	0.49	-0.60
KN3_6	Sweeps with mop	Sweeps with broom	Culturally appropriate picture introduced	0.65	-0.35
KN4_2	Stamp in wrong place	Theme unchanged	Culturally appropriate picture introduced	0.37	0.43
KN4_3	Balanced scales	Theme unchanged	Culturally appropriate picture introduced	0.67	0.48

Item No.	Original Item	Replaced Item	Reason	Difficulty Index (P)	Discrimination Power (D)
KN4_4	South and North America	Rearranged with (KN4_6) Wind in two directions	Difficult	0.15	0.27
KN4_5	Rooster on nest	Theme unchanged	Culturally appropriate clear picture introduced	0.36	0.43
QR2_4	5 birds	Theme unchanged	Replaced with local birds	1.00	0.64
QR4_2	Mathematical sequential order	Rearranged with QR4_3	Difficult where QR4_3 was easier	0.43	0.31
VS2_4	5 pieces placed correctly	Level rearranged level 3	Difficult (please see Table 9)	0.53	0.73
VS2_5	6 pieces placed correctly	Level rearranged level 3	Difficult (please see Table 9)	0.27	0.36
VS2_6	6 pieces placed correctly	Level rearranged level 3	Difficult (please see Table 9)	0.00	0.00

The above table suggests that some items (such as 7 were placed at 8 and original 8 were replaced at 7) were reorganized within the testlets. Besides these, pictorial presentations of some items (such as Bottle with straw, Sweeps with broom) were replaced by culturally appropriate pictures. These items were difficult because participants were unfamiliar with the uncommon items. However, due to the difficulty order of the original level, one level was re-arranged with another level (e.g. level 3 and 2 of VS2).

Table 8

Modified items based on P and D value (Verbal domain)

Item No.	Original Item	Replaced Item	Reason	Difficulty Index (P)	Discrimination Index (D)
KN_10	Child Drinking	Local Child Drinking	Culturally appropriate picture introduced	1.0	-0.21
KN_11	Cutting a piece of paper	more clear picture was introduced wit same theme	Culturally appropriate picture introduced	1.0	-0.21
KN_12	Boy Running	Local boy running	Culturally appropriate picture introduced	1.0	-0.21

Item No.	Original Item	Replaced Item	Reason	Difficulty Index (P)	Discrimination Index (D)
KN_13	Tying shoe	more clear picture introduced with same theme	Culturally appropriate picture introduced	1.0	-0.21
KN_14	Writing on a paper	Theme unchanged	Culturally appropriate picture introduced	1.0	-0.21
FR2_1	Cat and ball playing	Local Cat introduced	Culturally appropriate picture introduced	0.69	0.23
FR2_2	Laundry	Theme unchanged	Culturally appropriate picture introduced	0.69	-0.59
FR2_3	Puzzle	Theme unchanged	Culturally appropriate picture introduced	0.67	-0.64
QR2_4	2 Dogs	Theme unchanged	Culturally appropriate picture introduced	1.00	-0.47
FR3_3	Base ball	Cricket ball	Culturally appropriate picture introduced	0.90	0.04
FR3_4	Chop-stick	Fork	Culturally appropriate picture introduced	0.99	-0.01
VS3_1	In front of girl	Rearranged as VS3_3	Difficult where VS3_3 was easier P=(0.98)	0.66	0.22
VS3_2	Behind the girl	Rearranged VS3_4	Difficult where VS3_4 was easier P= (0.99)	0.54	0.11
FR4_3	Melted ice-bergs in Caribbean country	Melted ice-bergs in Bay of Bengal	Culturally appropriate theme introduced	0.42	0.41

The above table presents the adaptation that the study had made through item analysis. Usually, major adaptation was done by replacing the cultural friendly local item so that participant can express the theme accurately.

Table 9

P and D value of selected items of Visual Spatial Processing (level 3 and 2)

	Item1	Item2	Item3	Item4	Item5	Item6
Visual spatial processing level 3 original						
P	0.76	0.81	0.57			
D	-0.38	-0.48	0.08			
Visual spatial processing level 2 original						
P	1.00	0.87	0.60	0.53	0.27	0.00
D	0.64	0.64	0.82	0.73	0.36	0.00

Table 9 (above) represents the modifications that were done through the item analysis. Visual Spatial level 3 of Nonverbal domain was rearranged into level 2, as level 3 was correctly answered by majority whereas level 2 of Visual Spatial Processing was more difficult than the former .

Internal Reliability of Items

Besides, analyzing the P and D value, the researcher also computed the Cronbach's Alpha equation to justify the internal reliability of the items. (Cronbach, 1951). George and Mallery (2003) have provided the following rules of thumb for explaining the coefficient. A high value for Cronbach's alpha indicates good internal consistency of the items. It does not mean that the scale is uni-dimensional. Table10 (below) presents the internal constancy in the items of SB5 conducted in Bangladesh. This analysis was conducted among the 330 students for the items they responded. The overall reliability coefficient (α=0.84) suggests that there is high and increasing correlation among the items. Findings reveal that there is little bit lower alpha value for non verbal domain compared to verbal domain. The Cronbach's alpha coefficient for the all nonverbal items (items 152) is 0.82 suggesting that the items have relatively high internal consistency and for verbal items (items 141) is 0.75 indicating an acceptable internal consistency. This might be due to the difficulty in responding to verbal items where verbal expression is not given preference in the school examination marking system.

Table 10

Internal consistency (Cronbach's alpha) among items

Coefficient of Reliability	Average Inter item Covariance	Scale Reliability Coefficient
Non verbal (item-152)	0.0111548	0.82
Verbal (item-141)	0.0095924	0.75
Overall Items (item-293)	0.0076545	0.84
N	330	

Construction of Norm

The standardization, a research process of testing a normative group, was determined through the SB5 for the construction of norm. It consisted of a representative sample of 3300 students with approximately 300 subjects at each age group between the age ranges of 6 - 16 years. Based on the raw scores obtained from the ten subtests, age norm was calculated separately for the 11 age groups. Detailed process was discussed in chapter three (Method section). However, in a nut shell, in the first stage the raw scores obtained by the participants were the basis to calculate the IQ of an individual. After finding out the age specific mean and standard deviation, z scores (both small z and big Z) were calculated from them. Following the similar procedure, converting the raw scores into z score as a measure of standardization process, norm was developed from separate subtests (e.g., fluid reasoning, knowledge etc). These standardized scores were then reflected in the two domains of SB5 test i.e., non-verbal and verbal IQ, to capture both non-verbal and verbal ability. Finally the IQ was developed from the norms that were developed from the raw scores. This section, developing the norm and IQ, states and compares the sub test wise age specific descriptive statistics (mean, standard deviation) from the raw scores.

Table 11

Mean and SD of Raw Scores- All Students in Non verbal Domain

		Domain - Non Verbal				
Age		RS_NV_FR	RS_NV_KN	RS_NV_QR	RS_NV_VS	RS_NV_WM
6	Mean	11.5	10.6	14.4	14.1	18.4
	SD	1.4	1.7	1.8	3.0	3.7
	N	300	300	300	300	300
7	Mean	15.5	12.0	14.4	16.8	18.8
	SD	1.5	1.7	1.5	2.1	2.7
	N	300	300	300	300	300
8	Mean	16.0	12.2	14.7	17.3	19.4
	SD	1.0	1.8	1.7	2.0	2.7
	N	300	300	300	300	300
9	Mean	15.9	12.9	15.1	17.2	19.5
	SD	1.0	1.9	1.6	2.6	2.6
	N	300	300	300	300	300
10	Mean	16.1	12.9	15.1	17.9	20.1
	SD	0.9	1.7	1.3	1.3	2.6
	N	300	300	300	300	300
11	Mean	16.3	13.3	15.4	18.1	20.8
	SD	1.2	2.4	1.5	1.6	2.8
	N	300	300	300	300	300
12	Mean	16.4	13.3	15.7	18.4	21.0
	SD	1.3	2.3	1.5	1.6	2.8
	N	300	300	300	300	300
13	Mean	23.4	14.9	17.0	19.1	23.0
	SD	2.7	3.0	2.2	1.9	2.9
	N	300	300	300	300	300
14	Mean	24.3	16.0	18.1	19.7	24.2
	SD	2.7	3.5	2.6	2.1	3.2
	N	300	300	300	300	300
15	Mean	24.9	16.6	18.8	20.5	25.1
	SD	2.6	4.0	3.1	2.5	3.4
	N	300	300	300	300	300
16	Mean	25.5	18.0	20.2	21.3	26.7
	SD	2.5	4.0	3.3	2.8	3.4
	N	300	300	300	300	300

The data in the above Table 11 postulates the age specific descriptive statistics for every subtests of non verbal domain. The mean score of Fluid Reasoning of nonverbal domain ranges from 11.5 to 25.5, Knowledge from 10.6 to 18.0, Quantitative Reasoning from 14.4 to 20.2, Visual Spatial Processing from 14.1 to 21.3, Working Memory from 18.4 to 26.7 for ages six to sixteen years respectively. As age increases the mean score of sub test also increases. However, for Fluid Reasoning, there is a higher mean score for age 8 and 13 than the usual. And similar trend was found for other sub-tests. This might be due to test's format and the base points that they achieved during testing. The standard deviation ranges from 0.9 to 2.7 indicating there is a low variation with the scores obtained in each age group. However, the standard deviation is usually lower in low age groups showing better consistency in the scores.

Similarly, the following table describes the age specific descriptive statistics for every subtests of verbal domain. The mean score of Fluid Reasoning of verbal domain ranges from 10.3 to 15.8, Knowledge from 27.3 to 51.4, Quantitative Reasoning from 11.5 to 15.4, Visual Spatial Processing from 12.7 to 18.1, Working Memory from 10.5 to 19.2 for ages 6 to 16 years respectively. As age increases the mean score of sub test also increases. However, for Knowledge, there is a higher mean score for age 8 and 13 than the usual. And similar trend was found for other sub-tests. The variation is similar to the findings of nonverbal domain. The standard deviation ranges from 1.0 to 7.7 indicating there is a low variation with the scores obtained in each age group. However, the standard deviation is usually lower in low age groups showing better consistency in the scores.

Table 12

Mean and SD of Raw Scores- All Students in Verbal domain

	Domain - Verbal				
Age	**RS_V_FR**	**RS_V_KN**	**RS_V_QR**	**RS_V_VS**	**RS_V_WM**
6	10.3	27.3	11.5	12.7	10.5
	2.3	5.5	1.5	2.1	3.8
	300	300	300	300	300
7	10.6	28.9	11.6	13.0	10.6
	2.7	6.2	1.4	2.3	4.0
	300	300	300	300	300
8	11.1	30.4	11.8	13.4	11.5
	2.3	5.5	1.1	2.4	4.0
	300	300	300	300	300
9	11.5	30.8	11.9	13.7	11.5
	2.3	6.2	1.4	2.1	4.0
	300	300	300	300	300
10	13.4	40.9	12.1	14.4	12.3
	1.3	6.8	1.0	1.8	4.1
	300	300	300	300	300
11	13.9	42.5	12.4	14.9	13.4
	1.7	6.8	1.2	1.9	4.7
	300	300	300	300	300
12	14.0	43.8	12.5	15.1	13.1
	1.1	7.0	1.3	1.7	3.9
	300	300	300	300	300
13	14.3	46.9	12.7	15.7	14.0
	1.4	5.1	1.7	2.2	4.0
	300	300	300	300	300
14	14.8	47.9	13.6	16.5	15.8
	1.6	6.1	2.5	2.6	5.1
	300	300	300	300	300
15	15.4	49.0	14.2	17.1	17.0
	1.8	7.7	3.1	2.9	4.4
	300	300	300	300	300
16	15.8	51.4	15.4	18.1	19.2
	2.4	7.7	3.9	3.0	4.5
	300	300	300	300	300

The ranges of weighted mean (frequency weighted) for every subtest of non verbal and verbal domain were then calculated. The below result in Table 13 figured out the

weighted mean for Fluid Reasoning of nonverbal subtest ranges from 11.7 to 25.7, Knowledge ranges from 10.9 to 18.9, Quantitative Reasoning ranges from 14.6 to 20.8, Visual Spatial Processing ranges from 14.7 to 21.7, and Working Memory ranges from 19.2 to 27.1 for age 6 to 16 years respectively. It can be found that as age increases the weighted mean also increases. However, the standard deviation is almost 1.0 in lower age group showing more consistency between the answers and scores of the participants.

Table 13

Weighted Mean and SD of raw scores- All students (Non Verbal)

	Non verbal									
Age	RS_NV_FR		RS_NV_KN		RS_NV_QR		RS_NV_VS		RS_NV_WM	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
6	11.7	1.4	10.9	1.7	14.6	1.8	14.7	2.4	19.2	3.5
7	15.7	1.4	12.2	1.5	14.6	1.5	17.0	1.8	19.1	2.5
8	16.1	1.0	12.5	1.7	14.9	1.7	17.6	1.8	19.8	2.6
9	15.9	1.0	13.1	1.9	15.2	1.7	17.6	2.1	19.9	2.3
10	16.2	0.9	13.1	1.7	15.2	1.4	18.0	1.3	20.4	2.7
11	16.4	1.2	13.8	2.4	15.5	1.6	18.2	1.5	21.2	2.8
12	16.5	1.5	13.7	2.4	15.8	1.6	18.5	1.5	21.3	3.0
13	23.7	2.6	15.5	2.9	17.3	2.3	19.3	1.8	23.3	2.9
14	24.6	2.6	16.7	3.4	18.5	2.6	20.0	2.1	24.6	3.1
15	25.2	2.5	17.5	3.8	19.3	3.1	20.8	2.6	25.5	3.3
16	25.7	2.4	18.9	3.8	20.8	3.2	21.7	2.8	27.1	3.3

The result below (Table 14) shows that the weighted mean of raw scores varied by age in verbal domain as expected. The weighted mean of Fluid Reasoning subtest ranges from 10.8 to 16.2, Knowledge ranges from 28.5 to 52.5, Quantitative Reasoning ranges from 11.7 to 16.3, Visual Spatial Processing ranges from 13.1 to 18.6, and Working Memory ranges from 11.9 to 20.3 for ages six to sixteen years respectively. However, the result shows lower standard deviation for Fluid Reasoning and Quantitative Reasoning on the other hand,

standard deviation for Knowledge and Working Memory were higher. This was due to difficulty for the participants to memorize the topics consistently.

Table 14

Weighted Mean and SD of Raw Scores- All Students (Verbal)

	Verbal									
Age	RS_V_FR		RS_V_KN		RS_V_QR		RS_V_VS		RS_V_WM	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
6	10.8	1.8	28.5	4.8	11.7	1.4	13.1	2.1	11.9	4.1
7	11.3	2.1	30.2	5.1	11.8	1.5	13.4	2.3	12.1	4.3
8	11.6	1.9	31.4	4.8	11.9	1.1	13.9	2.5	12.9	4.3
9	12.0	1.9	32.0	5.3	12.0	1.2	14.0	2.2	12.9	4.4
10	13.5	1.2	42.0	5.8	12.2	0.9	14.6	1.8	13.7	4.2
11	14.1	1.8	43.6	5.8	12.5	1.4	15.2	2.0	15.0	5.0
12	14.1	1.2	44.9	5.9	12.6	1.5	15.3	1.7	14.3	3.8
13	14.4	1.4	47.4	4.6	13.0	1.9	16.0	2.3	15.4	4.1
14	15.0	1.6	48.7	5.6	14.0	2.9	16.9	2.7	17.4	5.2
15	15.6	1.8	50.2	6.9	14.9	3.6	17.6	3.0	18.1	4.4
16	16.2	2.6	52.5	7.1	16.3	4.3	18.6	3.1	20.3	4.5

The Z score was calculated through the mean and weighted mean from the raw scores of the subtest. Then, SAS was calculated separately for the two main domains - verbal and nonverbal from their sub tests. From this mean and SD the small z were calculated. Then following the equation 2, the big Z was calculated. This big Z was then ranked in 19 scores group (McGrew & Woodcock, 2001) and tabulated in each 10 subtests for each age level (Appendix11). The study depicts this as norm of urban students of Bangladesh for age group 6 to 16 years (Table 15). In the similar manner, the scores of each SAS was further computed and converted to Area SAS where the above formula was followed. The Area SAS was calculated to transform the scores into verbal and nonverbal IQ. The VIQ and NVIQ thus calculated following that process. The sum of the two Area SAS were again calculated following the above steps and finally to convert the two Area SAS into the third area to

obtain the Full Scale IQ. Standardized tests are also norm referenced; i.e., test scores are interpreted with reference to the scores. However, the IQ might not be bounded by age rather than it might depend on performance on the raw scores. Hence, a participant might receive higher IQ and again a participant from higher age might belong to low IQ based on his performance in the subtests. The above table summarizes the final IQ that was earned by the individuals through the raw scores from the subtests. The IQ, produced from the SB5 (adapted in Bangladesh) highlight that the mean IQ of Non-verbal, Verbal domain and Full Scale IQ. The Table 15 (below) states the IQ distribution by age group constructed from the norms developed from this sample.

Table 15

Range of IQ by Age Group

Age	Three IQs	Range Min-Max		Mean	SD(±)
6	NVIQ	72	107	91.54	5.45
	VIQ	63	89	75.75	3.83
	FSIQ	93	152	117.61	9.13
7	NVIQ	73	102	88.63	4.96
	VIQ	59	81	73.55	3.55
	FSIQ	92	130	111.60	6.46
8	NVIQ	68	84	79.88	2.68
	VIQ	59	74	66.26	3.81
	FSIQ	89	109	97.77	3.76
9	NVIQ	68	98	87.55	5.63
	VIQ	59	79	69.81	4.02
	FSIQ	88	127	107.22	7.09
10	NVIQ	79	95	89.25	2.13
	VIQ	65	82	73.95	2.39
	FSIQ	102	129	112.11	4.43
11	NVIQ	74	98	89.15	2.56
	VIQ	65	84	72.12	2.85
	FSIQ	93	130	110.61	4.99
12	NVIQ	75	99	89.24	4.24
	VIQ	59	76	69.31	3.23
	FSIQ	90	124	107.96	5.90
13	NVIQ	68	99	86.04	5.61
	VIQ	59	89	75.18	3.89
	FSIQ	87	145	111.04	8.26
14	NVIQ	68	99	85.76	6.26
	VIQ	64	87	74.25	3.34
	FSIQ	90	134	109.30	7.91
15	NVIQ	68	102	84.70	9.22
	VIQ	59	92	74.70	4.40
	FSIQ	88	148	109.52	11.38
16	NVIQ	68	105	86.65	8.87
	VIQ	67	89	75.54	3.35
	FSIQ	90	140	112.76	11.08

Table 16 suggests that the mean IQ of ages from 6 to 16 years varies by age groups. Findings reveal that the mean of full scale IQ had a sudden decrease at age 8 and 12. These might be because of inter-school migration during age 8 (probably class three) and during age 12 (probably class six). It is to be noted the group's performance is the basis for the tests norms.

Table 16

Correlation between Bangladesh and USA norm based on SB5

Variable	Mean	SD	Min	Max	Correlation
FSIQ_SB_BD	109.80	8.98	87	152	0.65
FSIQ_SB_US	76.70	12.99	46	146	
NVIQ_SB_BD	87.12	6.42	68	107	0.38
NVIQ_SB_US	85.24	14.43	42	152	
VIQ_SB_BD	72.73	4.58	59	92	0.48
VIQ_SB_US	70.64	12.72	43	139	

To see the consistency with the norm developed in Bangladesh through the current study, a correlation with the SB5 original was estimated. It states that there was 65 % correlation between the scores that the students obtained through the norm of original SB5 scale and the norm developed from the current study.

Table 17

Correlation between Male and Female within the Divisions

Variable	Mean	SD.	Min	Max	Correlation	Divisions
NVIQ_SB_BD_M	87.34	6.34	68	102	0.82	All
NVIQ_SB_BD_F	86.91	6.47	68	107		
VIQ_SB_BD_M	72.91	4.52	59	92	0.75	All
VIQ_SB_BD_F	72.54	4.64	59	89		
FSIQ_SB_BD_M_All	110.3	9.0	88	148	0.92	All
FSIQ_SB_BD_F_All	109.4	8.8	87	152		

As we looked into the male-female score composition and their IQ, we found that there was 75% correlation (Table 17) between male (1685) and female students (1685) of the sample based on the norm developed by the current study. However, we also found better correlation in Chittagong, Barisal and Dhaka divisions (Appendix 11) relation to the IQ scores obtained by the males and females students.

Based on the FSIQ from this study and the distribution of population, categorizations of IQ ranges were constructed. Following Roid (2007), the categorization emphasizes qualitative description of any IQ status of children 6 to 16 years in urban cities. The following Table 18 describes 7 categories of FSIQ based on Age norm of Bangladeshi children of 6 to 16 years.

Table 18

Categories of FSIQ based on Age norm (6 to 16 years)

Qualitative Description	Score level (Based on FSIQ)		Approximate Population
Significantly Below Average		≤86	
Moderately Below Average	87	94	6
Below Average	95	104	17
Average	105	115	50
Above Average	116	123	17
Moderately Above Average	124	127	6
Significantly above average	128	152	3

Determination of Reliability

Test retest reliability was computed to determine the reliability of SB5. The most common method for finding out reliability of test score is by repeating the identical test on a second occasion. The reliability coefficient in this case simply the correlation between the scores obtained by the same students on two administration of the test based on the norm developed from this study. Test-retest reliability shows the extent to which score on a test can be generalized over different occasions. The higher the reliability, the less susceptible the scores are to the random daily changes in the conditions of the subject or of the testing environment (Anastasi, 1992). Then Pearson's product moment formula (correlation) was quantified to determine test retest reliability. Test reliability is an element in test construction and test standardization and is the degree to which a measure consistently returns the same result when repeated under similar conditions (Cortina, 1993). Test-retest reliability was constructed on the scores obtained twice with the same instrument on the same individual with one week of time interval. A reliable measure should produce very similar scores both times following a high correlation (Lawler, 1978, Anastasi & Urbina, 1997). In this study,

three IQ scores of SB5 (NVIQ, VIQ, FSIQ) were computed and compared with the three IQ scores of the first and second administration on same individuals.

Table 19 presents the descriptive statistics of the raw scores. It depicts that the mean scores were higher working memory (short-term memory checking) and lowest for fluid reasoning (logical sequence, puzzle). However, there was no significant difference in the raw scores by the two administration of test indicating that the responses were more identical. Supporting to this the correlation coefficient suggests a high relationship. For fluid reasoning, a very high correlation (0.98) was found between the two administrations of test. Table 19 portrays that the correlation coefficients varies from 0.85 to 0.98 indicating strong correlation between the raw scores.

Table 19

Descriptive Statistics and Correlation among the Subtests of Nonverbal Domain

Subtests	M	SD	Min.	Max.	P value	Correlation	Total Items
FR_{1*}	19.14	4.86	8	31	0.344	0.98	36
FR_{2**}	19.10	4.75	8	32			
KN_1	14.59	4.21	6	27	0.905	0.90	30
KN_2	14.60	4.10	7	29			
QR_1	20.89	3.53	11	30	0.332	0.86	30
QR_2	20.79	3.49	7	30			
VSP_1	20.53	4.32	3	33	0.978	0.89	22
VSP_2	20.54	4.32	3	34			
WM_1	26.76	3.74	5	34	0.382	0.85	34
WM_2	26.89	3.37	17	34			

*1 denotes First administration and **2 denotes second administration of test SB5-BD

Table 20 shows the descriptive statistics of the verbal domain. The verbal domain represents the verbal ability of each individual from the test. It was found that the scores were higher for working memory subtest like the nonverbal domain. There was no significant difference in the raw scores by the two administration of test except quantitative reasoning (mathematical analogy). The reason might be the quantitative answers from the respondents' based on assumptions. However, High correlation (0.98) was found between the two administrations of test for knowledge (vocabulary, idea) subtest. However, the correlation coefficient ranges from 0.88 to 0.98 by the subtests.

Table 20

Descriptive Statistics and Correlation among the Subtests of Verbal Domain

Subtests	M	SD	Min.	Max.	P value	Correlation	Total Items
FR_{1*}	17.38	3.00	5	27	0.449	0.90	22
FR_{2**}	17.39	2.88	10	27			
KN_1	19.86	4.78	13	26	0.097	0.98	44
KN_2	19.76	4.77	13	26			
QR_1	17.89	3.92	11	29	0.020	0.94	30
QR_2	17.78	3.85	7	30			
VSP_1	17.07	3.33	10	27	0.364	0.88	30
VSP_2	17.03	3.23	9	29			
WM_1	20.29	4.95	5	30	0.148	0.91	15
WM_2	20.23	4.87	5	30			

*1 denotes First administration and **2 denotes second administration of test SB5-BD

Table 21 presents the descriptive statistic of the Intelligence Quotient (IQ) scores. The IQ was calculated from the students' raw scores following the standard process. The mean IQ for nonverbal, verbal and full scale was not significantly different. However the correlation coefficient suggests a strong correlation between the first and second administration of the test (from 0.72 to 0.76). The correlation of nonverbal IQ was lower compare to verbal IQ.

This might be because the national Bangla medium education system has less importance on verbal communication skills. Along with this, the socio-cultural system might act as barrier to express their verbal views spontaneously.

Table 21

Correlation of IQ Scores

IQ Scores	Mean	SD	Min.	Max.	P value	Correlation
NV IQ_{1*}	90.41	6.05	68	114	0.617	0.72
NV IQ_{2**}	90.58	5.91	76	114		
V IQ_1	73.91	4.72	59	97	0.757	0.76
V IQ_2	74.04	5.21	59	100		
FS IQ_1	114.29	9.37	90	152	0.800	0.75
FS IQ_2	114.47	9.25	95	152		

*1 denotes First administration and **2 denotes second administration of test SB5-BD

Determination of Validity

In psychology, validity has two distinct fields of application. The first involves test validity, a concept that has evolved with the field of psychometrics: "Validity refers to the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests" Anastasi & Urbina, (1997) describe. In this study we conducted content related validity and contrasted group validity (studies of special group) as well as SB5 (adapted in Bangla through this study) vs. WISC-R (Bangla version) under criterion related validity.

Content Validity.

As discussed in chapter three, a concrete and comprehensive opinion based on SB5 test adaptation including conceptual, methodological issues, items for producing adapted

instruments, translation and questions of ethics in cross-cultural contexts were taken by the professionals (Appendix 4). The decisions for item acceptance and other issues for adapted instrument were made in assistance with the guidelines recommendation from the International Test Commission (Guideline D1., ITC, 2001). Timing issues were also specifically addressed within the context of adapted test in their discussion.

Criterion Validity.

The study explored the correlation with the different domain scores of SB5 (adapted in Bangla through this study) and WISC-R (Bangla version) and Contrasted Group Validity to meet the criterion related validity.

SB5 (adapted from this study) vs. Wechsler Scales (WISC-R, Bangla Version).

In order to examine the criterion related validity SB5 and WISC-R were administered on the same participants. The study considered 90 students from three age groups (age 7, 11, and 14). The study found that the mean of IQ varies by age in considering SB5-BD (NVIQ, VIQ, FSIQ) and WISC-R (Verbal, Performance, Full Scale). Findings reveal from the descriptive statistics that there were significant similarities between the IQ scores obtained by the two tests.

Table 22

Mean and SD of IQ Scores of SB5 (Standardized from this study) and WISC-R (Bangla Version)

Variable	Observation	Mean	SD	Min	Max	P value
Performance (WISC-R)	90	108.41	17.89	16	138	0.001
NVIQ SB BD	90	90.21	5.92	76	105	
Verbal (WISC-R)	90	99.06	13.61	26	144	0.001
VIQ SB BD	90	73.26	4.39	59	91	
Full Scale (WISC-R)	90	106.61	11.03	71	164	0.001
FSIQ SB BD	90	113.30	8.81	95	145	

The Table 22 states lower correlation with the verbal (SB5) and verbal (WISC-R) IQ scores for three age groups. This correlation also figured out similar trend for nonverbal (SB5, adapted from this study) and performance (WISC-R) and FSIQ (SB5-BD) and FSIQ (WISC-R) and Full Scale (FSIQ-BD) IQ scores. Studies suggest that WISC-R is comparatively difficult as well as obsolete to the students whereas SB5 is usually user-friendly and latest edition. Besides, this might be because the adapted version WISC-R (1974) might need to revise as this one was standardized on 1980. However, the English version of WISC is currently updated in 2001.

Table 23

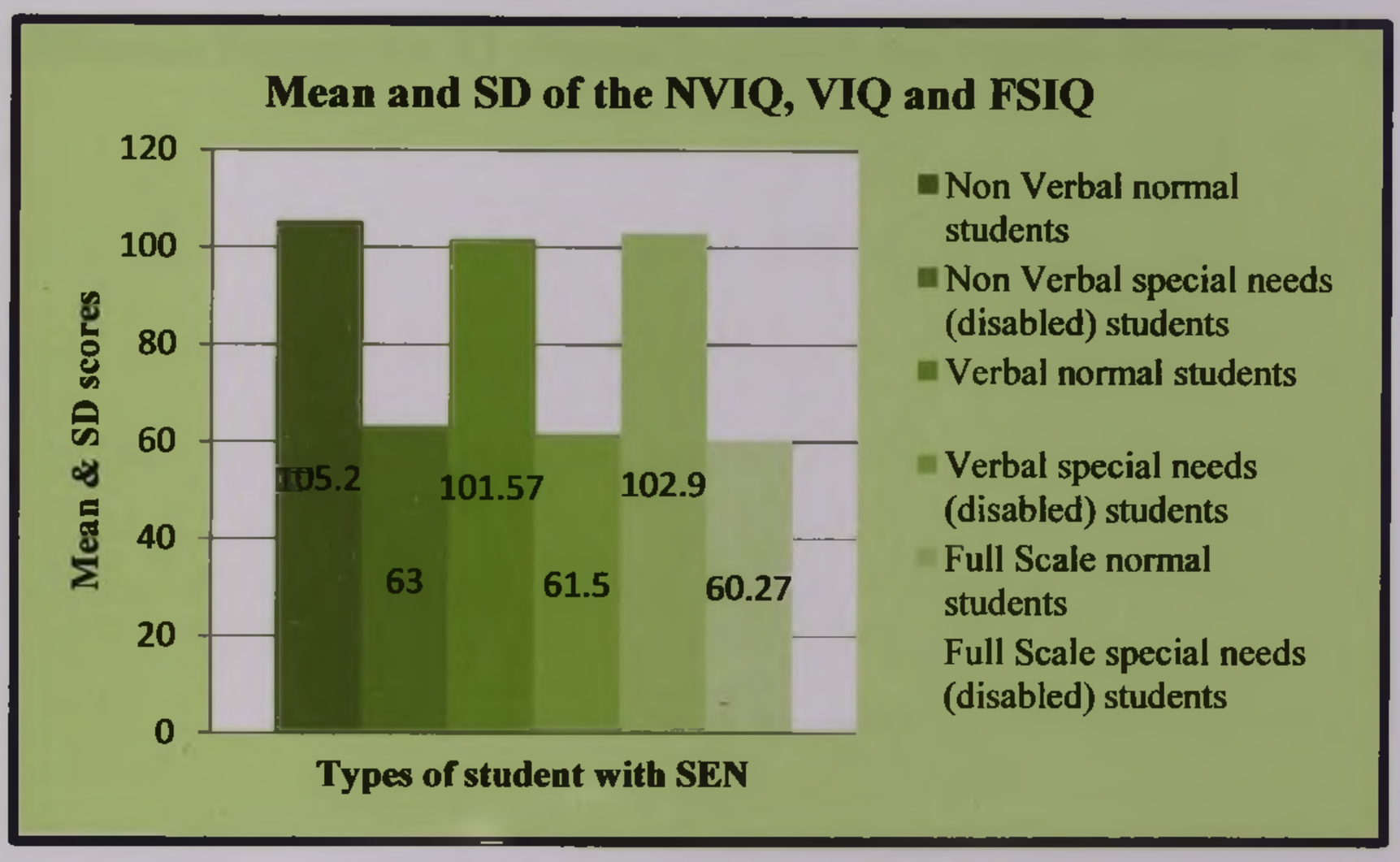
Correlation with scores of SB5 (Standardized from this study) and WISC-R (Bangla Version)

Age	NVIQ (SB5-BD) vs. Performance (WISC-R)	VIQ (SB5-BD) vs. Verbal (WISC-R)	FSIQ (SB5-BD) vs. FSIQ (WISC-R)
7	0.12	0.22	0.22
11	0.26	0.23	0.23
14	-0.09	0.13	0.13

Contrasted Group Validity / Studies of Special Group.

To find out the differences in IQ, the Stanford Binet Intelligence Scale Fifth Edition was administered on normal and different types of students with special needs (intellectually disabled, autism). With the three IQs [Non-verbal IQs (NVIQ), Verbal IQs (VIQ) and Full Scale IQs (FSIQ)] separate mean and standard deviations were calculated for the two groups. Figure 13 shows the mean and SD of the NVIQ, VIQ and FSIQ of the normal students and special needs student (following the norm developed by this study).

Figure 13. The Mean and SD of the three IQs of the Contrasted groups.



The result indicates that there is a similar trend among all the mean and standard deviations and the range is between 101 and 105 along with 4.16 and 5.32 respectively. Result indicates a low mean and standard deviation among special needs students.

Correlation of the Two Test Scores.

Table 24

't' (statistic) and P value between normal and students with special needs

Types of Sample	N	Mean	Standard Deviation	*t* value (statistic)	*P* value
Normal Students	220	102.9	±4.68	47.31	< 0.0001
Special Needs (Disabled) Students	220	60.27	±2.68		

To determine the validity between the two group't' value was also calculated to indicate the difference. The P value suggests that there is statistically significant (below 0.0001) difference between the IQ obtained by normal and students with special needs students.

CHAPTER FIVE

Discussion

Bangladesh has a dire need to improve and update the standard of existing assessment techniques in intelligence testing which are a continuous process and an integral part of educational instruction and development, as it determines whether or not the goals of education are being met. Assessment affects decisions about grades, placement, instructional needs and curriculum. To accomplish this goal, assessing individual's intellectual abilities is one of the prime requirements to ascertain one's potentialities. In broader perspective, the purpose of the present study was to improve and strengthen the existing standard of intelligence testing and upholds the trends by standardizing contemporary device for ensure the provision of identification, decision for educational placement and intervention services for the children with special needs in Bangladesh. In this context, standardization of a current intelligence test such as the Stanford Binet Intelligence Scale Fifth Edition (Roid, 2003) was required to serve this purpose.

At present the most useful evaluation method available for understanding of human intellectual abilities, and human knowledge is through standardized testing. However, validly and reliably used, standardized tests provide useful information to decision-makers that no other evaluation method can provide. Thus such tests have been developed and administered on a large scale population in advanced and high economic countries (Phelps, 2008). It is also true that, there is a necessity for a comprehensive approach to compare between individuals, as well as intra-individual performances. Considering that intelligence tests have long been regarded acceptable ways of predicting future outcomes, it occupies an important place in the educational and psychological landscape (Anastasi & Urbina, 1997). Likewise, Stanford-

Binet maintained a hybrid structure, combining point-scale and age-scale formats. This fifth edition also improved the psychometric characteristics of the test by introducing a parallel form and more representative norms from the earlier versions. The test would provide an estimate of the level at which an individual is functioning based on a combination of many different subtests or measures of skills (Becker, 2003).

This chapter provides a comprehensive description of the outcome of the study. It illustrates the findings of the standardization procedure with the various significant reflective features based on national and international perspective. Taking into consideration of the findings of this study, the core concepts - item analysis, construction of norm, reliability and validity of standardization process are discussed in this chapter. After the adaptation and completion of standardization process comprehensively and effectively through this study, the term SB5-BD was used in this chapter to discriminate from original SB5 and make the term reader friendly.

Standardization of SB5 in Bangladesh (SB5-BD)

Item Analysis.

In the context of psychometrics, item analysis is the procedure employed in test construction of qualitative evaluating each test item in terms of its content and form (Cortina, 1993). The item analysis is an important step in the development of any psychological tests. In this step statistical methods are used to identify test items that are too easy and/or too difficult. Item analysis is especially valuable in improving items which will be used again in later tests. But it can also be used to eliminate ambiguous or misleading items in a single test administration (Anastasi, 1997). The responses of students' performance in this test were similarly used to determine the difficulty and discrimination index for each test item. In this

study, hence, the higher this index value, the lower was the difficulty. The higher the discrimination index, the better the item could determine the difference, i.e., discriminate between those students with high test scores and those with low ones. The findings of the study in discriminating power and tending to centre on difficulty indexes of less than 50% were accepted for rearrangement, which also showed similar result in item analysis conducted by Huq in 1989 that the items have higher discriminating and difficulty indexes of 50% acceptance.

There is an increasing demand and need for psychological tests for different cultures and countries. In consequence, much greater awareness is being taken into consideration regarding the development or adaptation of test items. The items in the test must be culturally equivalent, where the meanings of the items need to be correctly translated and adapted so as to maintain the validity of the test in the new cultural context (Vijer, & Hambleton, 1996). Similarly, the findings of this study revealed that several items were adapted, modified or rearranged to keep the cultural fairness and sequential consistency from easy to difficulty order in comparison with age. In addition, by retaining the cultural uniqueness of the original SB5, the test was translated and adapted for Bangladesh culture. Further, Rodriguez and her associates in 2011 also emphasized in his study that there is a high correlation among intelligence, culture and age.

Adaptation of the items in Non verbal domain

Nonverbal domain presents a student's intellectual ability to analyze and solve problems without relying upon or being limited by language abilities. It intends to depict an individual's understanding for reorganization of visual sequences, analogies and causal relationships in illustrative and symbolic mode (Robinson, 2002).

Based on the findings, difficulty (Table 5, Figure 9, 10) and discrimination index (Table 6, Figure11, 12), the study adapted few items from original SB-5. Table 7, Table 8 and Appendix 10 provide the number of items modified, changed and retranslated during the adaptation process for both domains. By adapting the test to an examinee's functional level, the SB5 routing procedure increases the measurement precision by tailoring the difficulty of the items to the examinee's level of cognitive functioning. Vincent & Kamphaus (2011) conducted a study in USA for the construction and adaptation of nonverbal subtests of the Stanford-Binet Fourth Edition. This study recommended with a rationale for use of the adapted items in clinical and academic assessment practices.

However, in Routing- Nonverbal (Fluid Reasoning) subtest number of items that were found relatively difficult to understand was rearranged. Item 7, 26 and 28 (logical and mathematical problem solving) were replaced with the original item 8, 27 and 32 respectively. In Knowledge subtest, pictures that felt to be unfamiliar and uncommon to the students were modified. Items from level 2 (feeding a child), level 3 (drinks with straw from glass, eat with spoon from bowl, sweeps with mop), level 4 (stamp in wrong place, balanced scales, rooster on nest) were replaced with culturally appropriate pictures keeping the theme unchanged (Appendix 10). As to item 4 (South America and North America) was found difficult it was rearranged with item 6 (wind in two directions) of the same level. For Quantitative Reasoning, item from level 2 (5 birds) was replaced with picture of native birds (Appendix 10). Item 1 from level 4 (mathematical sequential order) was rearranged with the item 3 from the same level where item 3 was found easier in comparison to mathematical sequential order. In Visual Spatial, level 2 was reshuffled with level 3 in Nonverbal domain, as level 3 was correctly answered by majority whereas level 2 of Visual Spatial Processing was more difficult than the former . The study found no change or rearrange in the working

memory subtests. The Cronbach's alpha coefficient for the nonverbal items (items 152) is 0.82 suggesting that the items have relatively high internal consistency concluding that there is similar performance trend in the non verbal domain.

The items of the test (original SB5) were firstly edited and standardized for USA culture. The exclusive feature of the original SB5 sustains the nature of being culture free which required simply for the items to be rearranged and replaced in context to Bangladesh culture. Consequently, the themes of the items (people, common object) were consistent and unchanged instead of item elimination. So it has become apparent that the items of original SB5 thematically correlated to identify one's intellectual ability. Thus, in order to retain the original theme, the items were replaced with native content/symbol or object, made the item culture friendly, and often retranslated the question for better understanding of the students. However, as original SB5 is a test for 2 to 85 years of people whereas through this study, the *Bangla* version was standardized on students aged 6-16 years. Hence, this study did not attempt to modify the items of higher level (specially level 5 and 6), even though, the items were found too difficult. Moreover, when the items were found to be continuously difficult through difficulty index in several age groups, the items were adapted or modified considering the colour or the object or the language. The study by Huq, 1989 (previously discussed) also revealed that the adaptation of nonverbal subtest of Stanford Binet Intelligence Scale Fourth Edition (SB4) for use in Bangladesh was due to similar reasons, specially rearranging items and replacing local items. Besides, as adaptation is essential for standardizing a test, several studies suggest for the adaptation of Binet scale for use in their respective culture. Similarly, in India, Binet Scale has also been considered as a standard criterion for the assessment of intelligence. Among many adaptations, Kulshrestha (1971) adapted the Stanford Binet Intelligence Scale Form L - M in Hindi. Later, The Binet Kamat

Scale of intelligence (BKT) is also another Indian adaptation of the Stanford-Binet Scale of Intelligence. In this context, some of the test items and materials were replaced to suit Indian conditions, such as Indian coins, typically Indian pictorial scenes, vocabulary and Indian concepts. The intelligence scale assessed the child's skills in six areas: memory, language, conceptual thinking, reasoning, numerical reasoning, visual-motor coordination and social intelligence. A similar study on item fairness of nonverbal subtest of SB5 by Hurlow in 2011 suggests that there is little evidence of item bias between children and adolescents who are from a Latin country and Caucasian/White Non-Hispanic children with comparable ages, genders, and socioeconomic status. However, this study compares age specific items while the students were mostly representative of the urban middle class status.

Adaptation of the items in Verbal domain.

The verbal domain refers to the extent to which a student can approach words, sentences, written texts verbs, adjectives, as well as, the extent to which he/she can comprehend meanings, produce synonyms and antonyms, know the meaning and use of words, complete sentences with words omitted based on the word context and have a critical view towards written speech. Verbal skills may involve concepts as concrete or abstract ideas. It includes ability to analyze information and solve problems using language-based reasoning (Munoz –Sandoval, Cummins, Alvarado & Ruef, 1998). Verbal reasoning is important in most aspects of school work. Even the more abstract courses such as math and physics require verbal reasoning skills, as most concepts are introduced orally by the teacher. The verbal reasoning also reflects children's ability to explain verbal concepts clearly, provide rationale for their choices and explain conceptual information. Verbal ability, measured by the verbal IQ, is one of the most accurate predictors of academic success in formal school programs (Munoz –Sandoval et al., 1998).

Similar to the non verbal domain, adaptation was also done in the verbal domain. Likewise, 14 items were adapted from the original SB5. In Routing/Knowledge subtest, number of items that were found relatively difficult to understand were modified and adapted with culturally appropriate picture while keeping the sequence and theme unchanged for items 10 - 14. In Fluid Reasoning subtest, pictures that were felt to be unfamiliar and uncommon to the students were modified. Items from level 2 (cat and boy playing, laundry, puzzle), were replaced with culturally appropriate pictures keeping the theme unchanged. In level 3, baseball and chop- stick (utensil) were replaced with cricket ball and fork respectively (Appendix 10). In level 4, the item3 named as melted ice-bergs in Caribbean country were replaced with culturally appropriate themes (melted icebergs in Bay of Bengal). As the verbal analogy subtest is usually a complex concept for secondary age students with few exceptional correct responses. Finally, in level 5 and 6, the items had to be retained because of the higher – order age sequence. Since the age range of the scale starts from 2+ years the standard of the items of level 2 in Quantitative Reasoning subtest are suitable for very small children. Therefore, the items of level 2 were not taken into consideration for modification until further research on lower age levels. The picture of item 4 in level 2 was adapted keeping the theme unchanged to be culture friendly. In Visual Spatial Processing subtest, items 1 (In front of girl) and 2 (Behind the girl) of level 3 were found difficult and rearranged. The rearrangement was done by easier items 3 and 4 in place of 1 and 2. The study found no change or rearrange in the working memory subtests as it shown in nonverbal working memory subtests.

Findings reveal that there is a bit lower alpha value for verbal section compared to nonverbal section. The Cronbach's alpha coefficient for verbal items (items 141) is 0.75 indicating an acceptable internal consistency. This might be causes of our educational system

in Bangladesh where verbal communication is not given preference in the school grading system. Thus verbal items were difficult to response. Another current study evaluated the applicability of the Australian Adaptation of SB4 found considerably higher mean IQ scores than the normative US means (Rodriguez, Treacy, & Sowerby,1998) . Retaining the theme of original SB5, the findings of the study revealed that the items were significantly adapted, modified and rearranged through item analysis. The overall reliability coefficient (α= 0.84) suggests that there is high and increasing correlation among the items. The study concluded that the items were adapted for Bangladesh which was cultural friendly, changing the order, language through using the tools of item analysis. Since the item analysis in present study was the part of standardization of SB5 for use in Bangladesh, this criterion of standardization (item analysis) established a path for further accomplishment of norm development.

Construction of Norm of SB5-BD

Construction of age norm, as a part of standardization process, involves the administration of a test under uniform and standardized conditions to a large numbers of individuals at various ages. The two test areas for identifying ones intellectual ability are norm reference and criterion reference tests. Literature suggests that many standardized tests are norm referenced; i.e., test scores are interpreted with reference to the scores obtained from the sample. Norm-referenced tests are designed to examine individual performance in relation to the performance of a representative group. Criterion-referenced testing, unlike norm-referenced testing, uses an objective standard or achievement level. An individual is required to demonstrate ability at a particular level by performing tasks at that degree of difficulty. Scores on criterion-referenced tests indicate what individuals can do — not how they have scored in relation to the scores of particular groups of persons, as in norm-referenced tests. For the present study, age norm was computed from the raw scores obtained

by the subjects after administering the nonverbal and verbal subtests (five subtests from each domain) of the Stanford-Binet Intelligence Scale (Fifth Edition). In the study, the Binet Scale (SB5) was administered 3300 students from 11 age groups (300 from each age group) for the construction of norm. The original Stanford Binet Scale (Fifth Edition) was normed on a nationally representative sample of 4800 individuals, ages 2 to 85+ years. The sampling of original SB5 was matched on several variables (age, gender, ethnicity, geographic region and socio-economic levels) based according to national US Census, 2001 (Roid, 2003). In contrast, the present study maintained a standard technique in selecting an adequate number of sample sizes. In addition, the study focused on age, gender, metropolitan region and middle class status. However, the samples were purposively selected from 6 divisions whereas the seventh division was officially declared after the data collection (Appendix 5).

While in constructing the age specific norms, the study followed the standard procedure of calculation. It estimated the mean and SD of raw scores that was found from administering the test to individuals in assessing the IQ of two different domains (viz. non verbal and verbal) as well as the FSIQ. The FSIQ is computed as a sum of all the activities in the SB5; i.e., all subtests covering both the verbal and nonverbal domains of cognitive ability. Thus FSIQ is a global summary of the current general level of intellectual functioning as measured by SB5. In several times, researchers such as Carroll (1993) and Gustafson (1984) and Roid (2003) would describe the FSIQ as a measure of the hierarchical factor that exist among the scores of an intelligence test. The FSIQ score for the SB5 particularly strong in its predictive promise because it covers more factor than widely used batteries and includes a balanced coverage of both nonverbal and verbal aspects of each factor. FSIQ is intended to measure all possible aspects of intelligence that could occur across all cultures or settings. For example, some dimensions not represented in FSIQ include long term memory, auditory,

and kinesthetic abilities. Similarly, a study by Ken Richardson in 2002 suggests that IQ scores can be described in terms of sociocognitive-affective factors that differentially prepare individuals for the cognitive, affective and performance demands of the test. This study supports the findings of the present study for the construction of norm. The final outcome of this study is the qualitative description as well as categories of intellectual ability based on IQ scores. The qualitative descriptions and IQ ranges for age norm of 6 to 16 years children are Significantly Below Average ($\leq$ 86), Moderately Below Average (87-94), Below Average (95-104), Average (105-115), Above Average (116-123), Moderately above average (124-127), Significantly above average (128-152) and Above respectively . Similarly, the ability levels from the findings (Table 18) based on FSIQ Scores of SB5-BD are developed and followed according to the findings of the study of application of SB5 results to learning in the classroom by Carson & Roid (2004).

In Bangladesh, Huq (1992) developed the age norm of SB4 for interpretation of the raw scores. The partial composite standard age score (PCSAS) for the urban sample ranged from 32 to 166. This range of IQ also show a relationship to the present range of IQ which is from 86 to 152 (Table 18). Another similar study by the same author in 1980 for the construction of norm of WISC-R (1974) showed that the norm was calculated on basis of age. The IQ's ranged from 70 to 144 for the verbal scale, the performance scale IQ was from 66 to 143 while the Full Scale IQ ranged from 65 to 154. The IQ's obtained for three scales were quite similar to that of the original scale of the Wechsler Intelligence Scale for Children - Revised (1974).

The study found a strong correlation (r=0.67) between the FSIQ of original SB5 and SB5-BD (standardized through this study). In 1998, a study by Rodriguez and her associates also found a correlation between the IQ scores of Australian and Newzealand adaptations of

SB4 for Dunedin children. The research suggested that New Zealanders obtained scores at or above those of the US normative sample. The findings of the research stated that the Dunedin children obtained means comparable to the US norms at the significant level of 0.01.

Besides this, looking into the IQ scores between male and female participants the study found a stronger relationship in Dhaka (r=0.94) and Chittagong, Barisal (r=0.95) division based on the SB5-BD in comparison to the other three divisions. Moreover, considering the overall analysis of the scores represented by the 6 divisions, it can be induced that the descriptive statistics of their raw scores and IQ ranges are almost similar among male and female children (Appendix-11).

In this study male and female students scored almost similar score in their different subtests (Appendix 11) which may vary by age but not sex. This finding is also supported by several researches. Following this, Douglas and his associates (2006) identified similar result and depicted that there are no sex differences in overall general intelligence. According to their opinion, males are often observed to average higher scores on some tests of spatial ability, mathematical reasoning and targeting, while females are often found to average higher on some tests of memory, verbal ability, and motor coordination within personal space.

However, this study purposively selected the students from urban schools of divisional city which may result a higher correlation. As discussed earlier, a student of lower age might achieve higher FSIQ based on his intelligence shown in the raw score, and similarly, students of upper age might have lower FSIQ due to his poor intelligence in raw score. Cognitive strength, measured by the FSIQ can be used to improve areas of academic under achievement or cognitive deficit (Mc Grew & Hessler, 1995). For example, an

individual may obtain an average score in non-verbal and below average score in verbal score which indicates that s/he is comparatively advanced in non-verbal domain and need to improve verbal domain.

The intellectual ability of a child shows a gradual increase with age. This is because, as the child develops, s/he is exposed to more complex and new situations in life where he/she has to deal effectively and thus the child acquires new ideas and knowledge about the world around her/him. The present test results also showed a gradual increase of mean along with the increase of age (Table 11 & 12). Hence, FSIQ is not a constant issue rather along with subsequent counseling and guidance it may change or increase as age increases. Based on the IQ results, the theme of each test domain can be used to determine ones intellectual strength and weaknesses in particular area (Sternberg & Kaufman, 1998, 2000). For example, an individual might be excellent in working memory (insists on short term memory) but may be less competent in knowledge subtest (Appendix 11). As a result, it will be convenient for a teacher to guide students in a more constructive instructional strategy.

The Nonverbal IQ (based on the five nonverbal subtests of the scales) can be used for assessing individuals with communication disorders, hearing impairments or deafness, autism, specific learning disabilities, traumatic brain injury, and other conditions such as aphasia or stroke where linguistic ability is limited. The Verbal IQ (based on the five verbal subtests) can be used for special cases of orthopedic or visual impairment where emphasis is placed on oral presentation of verbal items (Roid, 2003). The SB5-BD can play a positive role in determining the standard of classroom teaching and learning process. Besides, the five factors comprising of the two domains (e.g. Nonverbal Fluid Reasoning, Verbal Knowledge etc) can help in identifying a student's strength and weakness (Fletcher, Foorman, Boudousque, Barnes, Schatschneider & Francis, 2002). Moreover, in several studies Carson

(2004), Clifford (2008) separately suggests that the thematic idea can be practiced in the classroom to reduce the weakness of students in a specific test area. The results eventually can guide teachers to prepare a student's Individualized Educational Plan.

Reliability of SB5-BD

The reliability of a test score refers to its precision in measuring the true attributes of a person and its consistency across sets of item, multiple testing occasions and other conditions that affect scores stability (Roid, 2003). The present study attempted to investigate the quantitative index of reliability for SB5-BD scores including test retest stability. The most common method for finding out reliability of test scores is by repeating the identical test on a second occasion, particularly after a week gap. The reliability coefficient in this case is simply the correlation between the scores obtained by the same persons on two times administration of the test.

The IQ scores (Nonverbal, Verbal and Full Scale) of the study were calculated with multiple formula (Pearson Product Moment Formula, Spearman's and Kendall's formula) to find out the reliability coefficients (Nunnaly, 1967, in Roid, 2003). The coefficients for the Non verbal (r=0.71), Verbal (0.76) and Full Scale IQ scores (0.75) of SB5-BD were with Pearson Product Moment. The study depicts that there was a consistency between the two correlations of tests scores. The subtest wise reliability ranges 0.88 to 0.98. Likewise, the original SB5 had a higher reliability of 0.90 to 0.93 (Roid, 2003). In a similar study, Madsen (1934) proved that the reliability coefficients of Stanford Binet IQs ranged from 0.65 to 0.94 respectively. Moreover, in Bangladesh, Huq (1992) determined the test retest reliability of the nonverbal subtests of SB4. The correlation thus computed was found to be 0.97 for urban sample. Measurement error is evaluated by examining the reliability of each test score. The

reliability of a test score refers to its precision in measuring the true attributes of a person and its consistency across sets of items, multiple testing occasions, and other conditions that affect score stability. Reliability for SB5 scores includes internal consistency, test-retest stability, and errors of measurement. Internal-consistency reliability ranged from 0.95 to 0.98 for IQ scores and from 0.90 to 0.92 for the five factor index scores. For the 10 subtests, average reliabilities (across age groups) ranged from 0.84 to 0.89, providing a strong basis for profile interpretation. Test-retest reliability studies were also conducted and showed the stability and consistency of SB5 scoring (Roid, 2003).

In another study, the reliability of the Stanford-Binet scale was determined by Lincoln, E. A. (2010). The findings of this research revealed that the correlation between the first and second examinations was 0.95. Test Retest reliability shows the extent to which scores on a test can be generalized over different occasions, the higher the reliability, the less susceptible the scores are to the random daily changes in the conditions of the subject or of the testing environment (Anastasi & Urbina,1997). In this study, means and standard deviations for test retest administrations of three IQ scores were consistent illustrating the stability of the scores obtained by students across time. Because there are many sources of random error across testing occasion, such as noise, distractions and moods of the students, the correlations are not expected to be as high as the internal consistency estimates presented in original SB5. Due to the effects of practice and familiarity with testing procedures, the mean scores of the test-retest sample may show some degree of improvement across administrations. Studies of test retest effects normally show that practice effects dwindle across intervals of several days or weeks (McArdle & Woodcock, 1997). Based on the SB5 studies and from comparisons with other IQ scale, the SB5 IQ scores appear to be quite stable and less affected by practice effects. (Gregory, 1996 cited in Roid, 2003).

Validity of SB5-BD

The most acceptable way of assessing an instrument's legitimate usefulness is through the use of validity studies (Anastasi & Urbina, 1997; Sattler, 1992). Validity studies, as outlined in the Standards for Educational and Psychological Testing are thought of as the "most important consideration in test evaluation" American Educational Research Association (AERA, 1999). In this current study (SB5-BD), validity is used to determine whether the assessments in question are in fact decent means of assessing intellectual abilities. Validity of test scores depends on the proper administration of the test by an experienced examiner and proper recognition of the unique characteristics of the individual examinee (Matarazzo, 1990 cited in Roid, 2003). Validity has several features and is established by the presentation of content related, criteria related and constructs related facts. Thus there is no single indicator of validity (Roid, 2003).

Expert opinions were taken into consideration while standardizing the test instruments. According to experts, the items and activities containing the SB5 were valid to be used for Bangladesh culture and content of the adapted items were highly correlated and consistent with the earlier Bangla version of SB4. On the other hand, Mark Pomplun and Michael Custer (2006) conducted a study on *the validity of* the measures of verbal and nonverbal working memory of SB5. The item mapping clearly demonstrated a parallel between increasing item difficulty and a progression of item characteristics that placed increasing demands on verbal and nonverbal working memory. The findings reveal that the higher correlations between SB5 verbal working memory and reading skills and between SB5 nonverbal memory and mathematics skills are consistent with past research.

Along with this content validity, criterion related validity and contrast validity were also experimented. There were significant difference ($p<0.001$) among the mean score between the special needs and normal students. Besides, there was lower correlation ($r=47.31$) between the two groups. This section of the study suggests that the test has ability to discriminate the normal and children with special needs. In 1991, Kline, Graham & Lachar investigated the contrast validity of nonverbal subtests of SB4 between students with verbal ability and students with reading problems. The findings depicted that there was significant difference and lower correlation among the scores of two groups that supports the outcome of the present study. Laurent et al. (1992) reviewed the Stanford-Binet Intelligence Scale-Fourth Edition and revealed the scale as a valid measure of general mental ability. The review also suggested that the SB4 could distinguish between groups of young students with differing intellectual abilities (e.g., mentally handicapped, gifted, neurologically impaired) and that the test correlated highly with scores on achievement tests. On the basis of validity information, recommendations for the use of the SB4 were made. A validation study by Tucker (1991) was conducted on the Stanford-Binet-Fourth Edition for using in the re-evaluation of learning-disabled students. The findings of the study ascertained that the SB4 scale is appropriate and effective assessment tool in evaluating strength and weakness of children with learning disability. In 2007, Abbott conducted a comparative Study of the Working Memory Scales of the WISC-IV and the SB5 in Referred Students. The study compared the working memory scales of the WISC-IV and the SB5 as both tests are used, in part, to develop academic interventions for students. There is a moderate correlation (0.6) between the two tests with 33 % of shared variance. The findings indicate that the two tests do not measure a similar ability and scores obtained on them should not be interpreted in the same manner. More research is needed to investigate the specific constructs measured and which test is most appropriate to assess working memory problems.

In another similar study, Askarian, Ali, Kambiz & Hassan (2011) computed the diagnostic validity for new edition of Tehran-Stanford-Binet Intelligence Scale in order to identify the children with learning disabilities. The results showed that this scale had the good diagnostic validity and desirable potential to identify students with learning disabilities. So according to them, this scale as a valid tool can be used for identifying students with learning disabilities can be used.

Apart from this, the study explored the correlation with the different domain from scores of SB5 and WISC-R Bengali version with a group of 90 homogeneous, nonexceptional school aged students, from three age groups (age 7,11 and 14). So, SB5-BD and WISC-R were administrated on the same students consecutively. Here, this study found a lower correlation 0.22, 0.23 and 0.13 between the IQs scores of two different intelligence scales (Table 23). Studies suggest that WISC-R (adapted in 1984) being outdated is comparatively difficult to the students where as original SB5 is usually standardized student-friendly and culture free test. Moreover, this lower correlation indicates that an individual's intellectual ability would never stagnant and permanent and should not assess through an obsolete and outdated intelligence test like WISC-R (adapted in 1984). Despite the limited nature of the sample, the fmdings suggest that the SB5 has a significant similar positive trends and relationship with the WISC-R. The tests displayed a moderate level of common variance. Inter co- relation partially supported the SB5 predictions of relationship between the two instruments.

Prior to this present standardization, the five nonverbal subtests of the Stanford Binet Intelligence Scale Fourth Edition (SB4) were also standardized in Bangladesh by Huq in 1991. The study validated the subtests with the three well known tests which were also

previously standardized for use in Bangladesh namely Independent Behavior Scale (IBS), Denver Developmental Screening Test (DDST) and WISC-R. The Pearson Product Moment coefficient of correlation was computed between the SB4 and three tests. Findings revealed that the correlation between SB4 and IBS was 0.71. Similarly, the correlation SB4 and WISC-R was 0.73; the two correlations seemed to be reasonably high. Finally the coefficient of correlation between SB4 and DDST was 0.57.

Another study that is adapted Bangla version of the WISC-R was standardized in Dhaka city by Huq (1980). The validity was computed to find out the separate correlation. The correlation between verbal subtests and school final examination was between 0.31 and 0.78. The correlation between performance subtests and annual examination ranged from 0.44 to 0.75. As for the full scale the correlation was between 0.06 to 0.69. Though the students or participant's school fmal examination record were not taken into consideration in this study for determining the correlation but all the participants were purposively selected on the basis of their school academic performance and that is average and above average students. Kush (2004) reviewed and described his study of comparison among the Stanford Binet Intelligence Scale Fifth Edition with the Wechsler Intelligence Scale for Children-Third Edition (WISC-III) (r = 0.84); and the Woodcock-Johnson III Tests of Cognitive Abilities (r = 0.78). The study found supplementary criterion-related validity between the two scales. In 1990, Hollinger & Baldwin examined the performance of 19 exceptional children on the Stanford-Binet Intelligence Scale, Fourth Edition (SB4) and the WISC-R. The results obtained for the naturally occurring sample of exceptional children indicate nonsignificant differences in performance between SB4 and WISC-R Full Scale IQ.

Similarly, the validity of the Stanford-Binet Intelligence Scale-Fourth Edition (SB-4) and that of the Kaufman Assessment Battery for Children (K-ABC) were investigated by

Emily and Robert (1987). In this study, the SB-LM was used as a criterion measure with which to compare SB4 and K-ABC scores. The study found significant correlation among the test scores. A paper prepared by Bivens in 1994 on Stanford Binet 4th Edition for adaptation in Australia. The correlation of the criterion validity in this study ranged between 0.67 and 0.83 which is statistically significant. Further, a longitudinal study of the Stanford-Binet and WISC-R with special education students by Covin, Theron & Sattler was conducted in 1985. Correlations between Stanford - Binet and WISC-R Full Scale IQs were significant in both ethnic groups, with r =0.60 for the total group.

Thus the processes of standardization are discussed based on study findings with several other significant literatures. Through various effective criterion measures, this process has been completed and finally it can be traced as a conclusion that the adapted standardized SB5-BD intelligence scale completed its standardization process through item analysis (modification and adaptation), construction of norm IQ range of children 6 to 16 years. Besides, SB5-BD has been established as a reliable tool through its consistency and accuracy of measuring intelligence of an individual. Moreover, standardization process also confirmed that SB5-BD is such a norm referenced standardized test that has been established by collecting relevant outline of evidence on individuals' intellectual ability so that educators can draw appropriate interpretations of assessment results and named as valid tool.

CHAPTER SIX

Implications

The author of the present study thinks that there are substantial outcomes from the research results on intelligence testing through the standardization of Stanford-Binet Scale Fifth Edition . A reliable and valid psychometric test device can be useful in educational and clinical evaluations. Its use has been and will continue to be influential in shaping educational policy and practice. In the past, scores from intelligence tests had led to wide spread ability grouping.

The most recently authorized 2004 version of IDEA (Individuals with Disabilities Education Improvement Act, P.L. 108–446) have emphasized critical guidelines that focused on the significance of this study for the identification, assessment and treatment of young children with special needs .With the contribution of legislation, the evolution of assessment process have focused from mere identification of children in need of early intervention programmes to increased emphasis on school-based services.

The results of these intelligence measures have at least three major areas of educational application.

Educational Provision

Since general intelligence plays an important role in many valued life outcomes, this research also suggests that IQ correlates with academic success eventually leading to future job performance and socioeconomic advancement (e.g., level of education, occupation and income).

The implications of SB5-BD provide a basis for ability performance relationships across major life arenas, including learning, work and daily life. The findings of this research suggests for making practical use of SB5-BD - for learning in classroom and classroom instructions. The outcome of the IQ scores can play a vital role in identifying the strength and weakness of low and high functioning students along with their special needs in classroom settings. The relative strength of nonverbal and verbal domain would be preferred for learning activities. For example, if the verbal domain is relatively stronger, the student is more likely to engage in learning through verbal means, such as reading, oral communication and through practice activities that emphasize the roles of speech and language. On the other hand, if the non verbal domain is relatively stronger, the student may be more likely to engage in learning activities that permit practice through nonverbal means.

The focal variables of this study were five factors (Knowledge, Fluid Reasoning, Quantitative Reasoning, Visual Spatial Processing and Working Memory) comprising of two domains (Nonverbal and Verbal), resulting in ten subtests (discussed in chapter two). The results unveil the pragmatic demands or recommendation of the above subtests for enhancing teaching learning strategies considering students learning style. The implications of each subtest in educational settings are stated below.

Nonverbal Fluid Reasoning (NFR).

A teacher can teach and evaluate a student's sequential and inductive reasoning ability through solving novel figural problems, sequences of pictured objects, geometric patterns in classroom situation while planning and preparing his / her regular lesson. At primary level, students can learn to match simple objects that are then placed in series (e.g. decreasing the size of the counting objects). Besides, students can learn to identify and extend the series of

sequential objects. At secondary level, teacher can teach students to continue a series of pictures to form repeating patterns (e.g. ball, bat, ball, bat etc.). Moreover, students can be taught by using the theme of this subtest through showing logical patterns of figural objects with one missing part. Research proves that it is a good measure of 'g' that assesses students' ability of correlational interpretation, attention span, perception of part versus whole, concentration and some degree of spatial analysis (Sattler, 1988).

Nonverbal Knowledge (NK).

To use the theme of this subtest in an educational program, a teacher can evaluate his/ her students' knowledge about common signals, actions, objects and the ability to identify absurd or missing details in pictorial material. At primary level, the activities, based on this subtest in a teacher's lesson plan, can measure a student's understanding of basic human activities (e.g. feeding a child, combing hair, clapping hands.etc) demonstrated in gestures. At secondary level, the activities will be more complex. The students can study pictures showing people in odd or inappropriate situations (e.g. girl with hair blowing in one direction while the wind blows the nearby trees in another direction) and point out the absurdity. The task requires students to have a basic level of common knowledge about people, nature and physical laws of the universe (Sattler, 1988). It also requires perception of detail, attention and concentration, inference, knowledge of science (e.g. how a balance works), and geography (missing nations on a world map). Teachers or educators also plan the activities to explain the absurdity vocally despite the presence of a visual illustration to assist the students. The students can point to the location and use gestures in addition to vocal speech to explain the silliness.

Nonverbal Quantitative Reasoning (NQR).

Based on this subtest, the educators as well as teachers can utilize their knowledge in their educational planning. They can judge students' ability to solve increasingly difficult prematheматic, arithmetic, algebraic, functional concepts and relationships depicted in illustrations. At primary level, the activities design on this subtest can measure basic concepts (e.g. bigger/ smaller), counting, addition using objects and pictures and recognition of numbers. At secondary level, students can be taught and assessed with increasingly complex activities with illustrations depicting figural series, functional relationships, linear transformations and logic or algebraic relationships. Research shows that VQR and VK subtests will have greatest relevance for school based learning and possible academic interventions (McGrew, Keith, Flanagan & Vanderwood, 1997).

Nonverbal Visual Spatial Processing (NVSP).

By using this subtest in the lesson, the teachers can plan to teach and assess student's ability for visualization and solution of spatial and figural problems. Hence, at elementary level, this subtest can be used in assembling puzzle like pieces and visual matching activities. At the secondary level, this subtest is to be considered as a unique and interesting as well as new challenging task to above average and higher functioning students. In this context, the students can duplicate familiar patterns such animals, objects (e.g. house, boats) and people in motion by properly arranging the object pieces.

Nonverbal Working Memory (NWM).

A teacher will be able to measure fundamental short term memory of her students with observable objects and utilize this skill in tapping sequential activities. The teacher must

plan his/her teaching activities containing this subtest with recalling a sequence of block taps. According to Reid, Hresko & Swanson (1996), students can learn to memorize and sort out the activities as well as chunk of numbers that are stored in short term memory.

Verbal Fluid Reasoning (VFR).

The activities in this subtest include early reasoning, verbal absurdities and analogies. Based on this, a teacher can judge the ability to analyze and explain, using deductive and inductive reasoning and problems involving cause - effect connections in pictures, classification of objects, absurd statements and interrelationships among words. In this context, the activities at primary level will require the students to verbally describe implied connections in pictured events, sorting and classifying pictured objects. The activities of secondary level are to be designed to assess verbal reasoning and completion of analogies such as "_ is to B as C is to _." Moreover, students of this level can learn by using their verbal abilities such as verbal fluency, vocabulary meanings and variations and solve problems such as guessing and checking (Carroll, 1993).

Verbal Knowledge (VK).

This test is termed as Vocabulary which measures general and crystallized ability and is applied in psycho educational settings. To judge the ability, the teacher can emphasize on students knowledge and memorization of concepts and language, and to identify and define increasingly difficult words. Likewise, this subtest can be used in elementary level in identification of body parts (on the students own body and on the child picture), toy objects and picture vocabulary. Besides, at secondary level, students are presumably influenced more by the effects of schooling and extensive reading. They can be more competent in learning more upper level vocabulary activities lead to higher literacy level, exposure to higher levels

of spoken and printed Bangla and English language. This subtest does not emphasize articulation but it requires an individual's ability to understand and explain the meaning of words.

Verbal Quantitative Reasoning (VQR).

The lower levels activities of this subtest are to be designed to measure counting of toys, basic addition and subtraction using pictured objects and/ or word problems. At secondary levels activities, students are to be taught in measurements, geometric and word problems with multiple methods of solution. As stated earlier, since this subtest is based on academic learning, a teacher motivates students' reading ability to solve increasingly difficult mathematical task involving the above outlined activities.

Verbal Visual Spatial Processing (VVSP).

Using this subtest, a teacher can determine the requirement of the ability to identify common objects and pictures using common visual spatial directions, indicating direction and position in relation to a reference point. For example, the activities in elementary level are to be designed containing with pictorial tasks requiring understanding of basic spatial concepts such as "behind" or "away from". At secondary level, more expressive language will be required to explain spatial orientations and directions in increasingly complex tasks.

Verbal Working Memory (VWM).

This subtest uses the activities of memory for sentences and last word. Through this test, a student requires the ability to demonstrate short term and working memory for words and sentences and to store, sort and recall verbal information in short term memory. To teach

students, the teacher can read short phrases and sentences aloud to the students, who then recall them accurately. The teacher can also ask sets of questions and the students recall the last word in each question. The students are required to answer each question, 'yes' or 'no'. There are many situations in the classroom where a student must selectively attend to portions of a teacher's messages. According to Roid, 2003, such efforts entail filtering out noise from other students to hear the important messages from the teacher. A low level of ability in selective listening would seem to be predictive of an individual's underachievement in the group – instruction methods of modern education. This subtest will provide information about students' cognitive deficits to teachers and parents.

Besides, a teacher will acquire self-confidence in handling students with diversity. This approach will promote peer interaction as well as create student friendly teaching learning classroom environment. The five factors of SB-BD has been proved (according to findings) as the precursors for the measurement of literacy skills (reading and mathematics). Educators will get a clearer picture of potential academic difficulties and determining which educational interventions may be helpful at the school level (Coleman, Buysse, & Nietzel, 2006). The implications of SB5-BD can also be considered in planning curriculum modification for children with special needs. The study reveals that the IQ scores will guide a teacher to have in-depth knowledge of a student's potentialities. Further, based on the findings and in relation to the qualitative categories as well as FSIQ Scores (Table 18), the author recommends the following instructional strategies that can be utilized while teaching children with special needs along with other students in mainstream or inclusive settings (Table 25).

Table 25

Relations between FSIQ Scores of SB5-BD and Recommended Methods of Instruction

Ability Level and FSIQ	Optimal Method of Instruction
Significantly Below Average (≤ 86)	Ensure that learning is at an appropriate slow speed, simple and supervised.
Moderately Below Average (87-94)	Provide very direct, hands-on-instruction.
Below Average (95-104)	At lower range, may benefit from plenty of direct supervision.
Average (105-115)	Students can thrive in learning in a traditional classroom format, with mixed
Above Average (116-123)	Can more readily acquire skills in collecting and gathering their own information.
Moderately Above Average (124-127)	Create opportunities for these individuals to seek and find their own information and provide information as needed, particularly in these information search skills.
Significantly Above Average (128-152) and Above	These individuals may enjoy reasoning things through on their own. Use more direct methods as needed, but remember that traditional classroom teaching methods may become boring for these students.

The above recommendations for optimal methods of instruction in classroom application based on FSIQ Scores of SB5-BD are similar and followed the study of application of SB5 results to learning in the classroom by Carson & Roid (2004).

In addition, the SB5-BD factors may show a difference among students with high and low abilities. Students with high ability will appear to be creative and rely on reasoning skills by which they can reach to decisions with confidence. They can be judged several tasks at a time. High functioning students require enriched environments to gear up their creativity and potentialities. On the other hand, students with low abilities will tend to avoid unstructured problems, become frustrated with too much demands from others, tendency to avoid mathematical problem solving issues, finds difficult to visualize problems through imagination, they are viewed as distracted, forgetful and inattentive by others. Moreover,

students with low functioning must have appropriate and simplified curriculum and teaching method with sufficient hands on instruction and plenty of direct supervision. Thus the above characteristics identified for this group has to be taken into consideration in the classroom teaching learning situation along with SB5-BD. The researchers think, incorporating the verbal domain in national curriculum would benefit the students to improve their verbal communication skills. Based on the rationale of the study (sketch out in chapter one), application and implication of all the ten subtests of SB5 are discussed in this chapter. It is to be mentioned that individuals' potentialities and abilities should be judged by using these subtests but not biased by the subjective opinion of the teachers and parents.

Screening, Diagnosis and Remedial Planning

Early detection and diagnosis is the single most important key for effective way to reduce the risk of developing secondary problems as well as availing reasonable preventive measures.

However, intelligence testing is the estimation of a student's current intellectual functioning through performance of various tasks designed to assess different types of abilities. The test scores will provide important information on how children's ability can be properly interpreted to help educators for developing appropriate educational strategies for remedial planning and intervention program and decisions for placement. Besides, information from tests is more scientifically consistent than from a clinical interview, as well as for legal matters, when decisions have to be made for disability issues, the standardized information from tests scores will help to overcome the personal judgment of the authority.

It is well established that since intelligence tests can reasonably predict levels of achievement, SB5-BD can be considered as a tool for identifying low, average and high intellectual functioning of students along with its relevant assessment techniques.

Accountability, Research and Evaluation

IQ measures are often included among outcome measures related to programme effectiveness. In addition, they are among the measures used in research to account for pupil characteristics. Besides, Stanford-Binet intelligence test is one of the important tools of psychological assessment used by professionals, psychologists in a variety of settings such as private offices, public and private schools, private and public mental health clinics and institutions, hospitals, the personnel offices of industrial companies and the counseling centers of colleges or universities are among others. This research states that the implication of IQ scores can play a major role in academic settings to determine developmental disability, prevalence of intellectual disability and other exceptional children with high or low capabilities. Additionally, there are other educational indications, such as eligibility criteria for service delivery and school accommodations.

It has become apparent that the implications of SB5-BD for future intelligence testing and for education are numerous. Assessment of intellectual qualities should go much beyond present standard intelligence tests, which seriously neglect important abilities that contribute to problem-solving and creative performance in general. Educational philosophy, curriculum-building, teaching procedures and examination methods should all be improved by giving attention to the structure of intellect as the basic frame of reference. The standardization of latest intelligence scale and its application in Bangladesh is not a new trend; rather it is a useful, dynamic and constant continuing practice in the field of assessment. Along with

above significant and detailed discussions throughout the study, it has become crystal clear that Stanford Binet is the exclusively psychometric and contemporary intelligence scale among its counterpart which have been standardized in various editions and adapted in several cultures only for the necessity of assessing intellectual ability of an individual and for the provision of intervention and remedial services for children with special needs. It can be concluded that the special features of SB5-BD also proves that various intellectual abilities in individuals can be improved by utilizing this standard assessment scale which holds a platform on top of all intelligence testing and goes beyond all the debate of traditional testing.

Recommendations

1. Classroom teaching –learning strategies can be modified by utilizing the theme of ten subtests of SB5-BD.
2. This study highly recommends utilizing the test in clinical setting for the identification and educational placement of children with special needs.
3. Further the outcome of the study can be executed in developing functional and simplified curriculum for children with special needs.
4. Educators as well as professional of the educational institutions can utilize this SB5-BD for the evaluation of their student's intellectual performance.
5. Government should take necessary initiative to utilize the ten subtests of SB5-BD for upcoming NCTB curriculum modification.

Recommendations for Further Research

1. The need for a holistic concept of a student's intellectual ability and for considering other age levels, further research are recommended along with an extended age range.

2. Deeming the present study a success, there needs to be further research taking into consideration students from rural and other diversified region to obtain a more comprehensive knowledge and understanding of the students' ability.

3. In order to further validate these results, this study or similar studies need to be replicated with a larger sample size to indicate a more significant difference. In addition, a more representative study is needed that would include more diverse subjects to generalize the population.

4. More research needs to be conducted regarding IQ testing with persons who have Intellectual Disability (ID). We need further data to determine whether the Stanford-Binet produce similar or different results. Furthermore, other IQ scales also need to be compared with the Stanford-Binet for persons who have ID. This is an important issue that needs to have careful review and study as the public policy implications are huge.

5. The Government may afford funds for establishment of Assessment and Counseling Unit in the department of Special Education within IER, DU and may also plan to set up a test taker training center in this unit to increase the human resource in the country so that the access of students with special needs in mainstream schools can be ensured.

Limitations

Since the standardization was limited to age (6 - 16 years) and for the sake of assessing children, the ten subtests (nonverbal and verbal) of the scale were taken into consideration for adaptation and standardization. The first limitation of the study was that the sample had been taken from a relatively small geographic region. Participants were considered from urban and metropolitan six cities area in Bangladesh. This means that the

results of the study might not generalize to children from other geographic regions such as rural area. Although the study would attempt to contain demographic variables representative of a national sample , participants might be limited in other variables such as economic status and ethnicity. Since the study would be comprised of only students between the ages of 6-16 years and results might not generalize to other age groups of original SB5. Finally, obtaining a sample size of 4400 participants, the study limits the applicability of the results to more global populations. Caution should be used when making assumptions regarding large populations utilizing limited sample sizes.

Over and above, the study had to regard other limitations in context to the desired time frame for test administration as given below:

- The authorities of few schools were less enthusiastic in pursuing the task of test administration that made the work lengthy.
- Considering the schools' academic formalities (e.g. class test, half yearly and final examination), test administration had to be deferred as well as in many situations, thus the pre-set schedule could not be achieved.
- Besides, the other major limitation at the field level was the school holidays (following the school calendar) which also acted as a serious back lock for the smoothness of the overall research work.
- Further, as test administration follows a systematic and standard procedure as given in the manual, the researcher had to obtain a thorough understanding of the test procedure. This required ample time to gain the proficiency in administering an individual test.

- Since it was an individual test, the researcher had to consider a student's pace of response.
- Finally due to lack of expertise on test administration, along with research assistant, the researcher conducted the test individual handed.

Conclusion

The research relevant to theory and practice in intelligence shows that the field of testing intelligence is active and dynamic. Also, it should be evident that intelligence researchers of the 21st century are addressing a broader, more complete concept of intelligence than was evident in the previous century. As related research in biology of the mind, emotion, neuropsychology, family dynamics and cognitive processing progresses to new findings, these results will be incorporated into increasingly useful models and theories of the workings of intelligence and how to assess individual intelligence. After completion of this study, the SB5-BD appears to be an effective measure of general intelligence across a specific age range (6 to 16 years) and the standardization sample appears to be a close match to the population on key demographic variables. IQ scores also cover a wide range of ability from the lower levels of moderate intellectual disability to the higher levels of intellectual giftedness (Table 18). As such, they will be helpful in assessing students with intellectual disability, learning disabilities and intellectual giftedness and interpretation of the global full scale IQ appears to have strong empirical support. Therefore, the present standardized test SB5-BD could be considered a unique achievement in the field of testing for a country like Bangladesh. The Stanford-Binet has been standardized, translated and adapted in many languages and used in many developed and developing countries. The necessity for a standard assessment scale was required to diagnose children with special needs for their

appropriate intervention and the outcome was the standardization of Stanford - Binet Intelligence Scale (Fifth Edition) in Bangladesh, SB5-BD.

Though there are constitutional, legislative and policy bindings and Bangladesh ratified the CRC and signed the entire International and Regional declaration on Education (reviewed in chapter two), the Government of Bangladesh has not yet undertaken significant steps to ensure education for children with special needs. Bangladesh is far behind in developing an effective education system for children with disabilities, and whereas the Government of Bangladesh established a special and integrated education system and NGOs are implementing special and inclusive education system as well. The educational programme of children with special needs remains under the Ministry of Social Welfare, which indicates that the educational issue of these children is being considered as a welfare concern, rather than a developmental subject. So, there is a big gap in incorporating children with special needs into mainstream education. In Bangladesh, the Education Policy provides provision for 'Education for All' and primary education is compulsory and free. Children with special needs are left out of this programme as their educational provision is seen as a welfare and charity issue. Under PEDP-II (Primary Education Development Project- II), it has been specified that in primary schools children with special needs with mild delays would be enrolled, but unfortunately this does not happen in practice. (Choudhuri et al., 2005).

With the overall discussion and observation, it is clear and apparent that professionals like school psychologists, educators are to be concerned while selecting an effective assessment scale to measure a student's intellectual and achievement ability. Based on their practical and realistic judgment and through common consent it has been proved worldwide that the Binet scales along with its rich tradition and popularity, solely occupies the field of assessment. Similarly, in compliance with the above view, the present study also standardized the latest revision of the Stanford Binet Intelligence Scale for use in Bangladesh. In addition there will

be a great impact as well as change on our education system by utilizing the test scores. On the other hand, it should be added that intelligence testing should not be conducted in a vacuum. Furthermore, by using previously mentioned norm-referenced tests, an assessment should include a variety of other data from a multitude of informants. Sattler (2001) suggests that norm-referenced testing should be accompanied by interviews with a parent, teacher and student; observations of the student during both the formal testing and natural environment (e.g. classroom, lunchroom, playground); and informal assessment procedures (e.g., district-wide criterion-referenced tests, school records). Such an assessment will provide the most accurate information by which educators can most effectively serve the student.

Though Government of Bangladesh has frantically taken several remarkable attempts to fill up all the loopholes in educational development but their effort will go in vein if the foundation of educational development such as intelligence testing would not be considered and included as main concern issue. Thus, it can be concluded that through this study the researcher would like to draw the attention of the Government and policy makers to be acquainted with the importance and the impact of assessment. Further, to include the need of assessment as a prerequisite criterion to justify a student's academic progress based on their potentialities. In this respect the Government can include the compulsion of assessment in education policy to provide effective guidance and counseling programme in all educational institutions. The Government should mobilize funds for test development and address the issue of curriculum modification to ensure the access of students with special needs in mainstream schools to fulfill the commitment of Education for All.

References

Abbott, E. N. (2007). *Comparative study of the Working Memory Scales of the WISC-IV and SB5 in referred students. Proquest Dissertations and Theses. Section 0817, Part 0525* 31 pages; [Ed.S. dissertation].United States -- West Virginia: Marshall University; Publication Number: AAT 1448545.

Ahmann, J.S., & Glock, M. (1981). *Evaluating student progress: principles of tests and Measurements, (6th ed.).* Boston : Allyn and Bacon, Inc.

Algozzine, B., Eaves, R., Mann, L., & Vance, R. H. *(*1993*). Slosson Full-Range Intelligence Test. New York: Slosson Educational Publications.*

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME) (1999).*Standards for Educational and Psychological Testing (3rd Ed).* Washington, DC: American Educational Research Association.

American Psychiatric Association, (2000). *Diagnostic and Statistical Manual of Mental Disorders* [DSM-IV-TR, 2000] (4th Ed.). Arlington, Va.: American Psychiatric Pub.

Anastasi, A. (1976). *Psychological Testing(4th Ed)*. Waterloo, AL, USA: McMillan.

Anastasi, A. (1992). *Psychological Testing* (4th Ed.). Singapore: Delhi Pearson Education.

Anastasi, A. (1992). What counselors should know about the use and interpretation of psychological tests? *Journal of Counseling and Development, 70 (5),* 610-615.

Anastasi, A., & Urbina, S. (1997). *Psychological Testing* (7th Ed.). NJ: Prentice-Hall.

APA. (1992). Ethical principles of psychologists and code of conduct. *American Psychologist, 47,* 1597—1611.

APA, (2000). *Report of the Task Force on Test User Qualifications.* Washington, DC: Author.

Armour-Thomas, E., & Gopaul-McNichol, S. (1998*). Assessing intelligence*. CA: Sage.

Askarian, M., Ali, A.G., Kambiz, K., & Hassan, P.S (2011*) The Diagnostic Validity For New Edition of Tehran-Stanford-Binet Intelligence Scale in Order to Identify the Children with Learning Disabilities. Australian Journal of Basic and Applied Sciences, 5(12)*: 1023-1029, ISSN 1991-8178.

Becker, K.A (2003). History of the Stanford-Binet Intelligence scales: Content and psychometrics. *Stanford-Binet Intelligence Scales, Fifth Edition Assessment Service Bulletin No. 1.*

Binet, A., & Henri, V. (1895). La psychologie individuelle. *L'Annee Psychologique, 2,* 411–465.

Binet, A. (1903). *L 'Etude experimentale de I' intelligence.* Paris: Schleicher Frères.

Binet, A., & Simon, T. (1905). Methodes nouvelles pour le diagnostic du niveau intellectual des anormaux. *L'Annee Psychologique, 11*, 191–244.

Binet, A., & Simon, T. (1908). Le developpement de l'intelligence chez les enfants (The development of intelligence in children). *L'Annee Psychologique, 14,* 1–94.

Binet, A. (1911). Nouvelle recherches sur la mesure du niveau intellectuel chez les enfants d'école. *L'Annee Psychologique, 17*, 145–210.No found

Binet, A. & Simon, T. (1916). *The development of intelligence in children: The Binet–Simon Scale*. Publications of the Training School at Vineland New Jersey Department of Research No. 11. E. S. Kite (Trans.). Baltimore: Williams & Wilkins. Retrieved 18 July 2010.

Bivens, M. (1994). *Critical Review of Stanford Binet* (4th Ed), Australia: ACER.

Bouchard, T. J., & Segal, N. L. (1985). *Environment and IQ.* NY: John Wiley.

Bracken, B. A. (2007). Observing preschool assessment-related behavior. In B. Bracken and R. Nagle (Ed.), *Psychoeducational assessment of preschool children, Third Edition* (pp.137-154). Lawrence Erlbaum Associates.

Cahan, S., & N. Cohen. (1989). Age versus schooling effects on intelligence development. *Child Development* 60:1239-1249.

Campbell, K. (2006). *Correlations between the WISC-IV, SB: V, and the WJ-III tests of achievement.* Proquest Dissertations And Theses 2006. Section 0817, Part 0288 17 pages; [Ed.S. dissertation].Marshall University. United States -- West Virginia: United States. Publication Number: AAT 1434477.

Carmines, E. G., & Zeller, R. A. (1991). *Reliability and validity assessment*. Thousand Oaks ,CA: Sage.

Carroll, J. B. (1993*). Human cognitive abilities: A survey of analytic studies.* NY: Cambridge University Press.

Carson, A.D. (2004). *Application of SB5 results to learning in the classroom*. Riverside Publishing Company.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*, 1—22.

Ceci, S. J. (1996). *On Intelligence: A Bioecological Treatise on Intellectual Development*. Expanded edition. Cambridge, Mass: Harvard University Press.

Chang, Y. (2008). Cross cultural assessment: A call for test adaptation (summer ed.) New stones: Association for Assessment in Counseling.

Charles S. (1996). *Major psychological assessment instruments*. Newmark (ed.) Boston : Allyn and Bacon.

Chase, D. (2005). *Underlying factor structures of the Stanford-Binet intelligence scales fifth edition* (Doctoral dissertation).

Chen J, Hale S, & Myerson J. (2007) *Predicting the size of individual and group differences on speeded cognitive tasks. Psychonomic Bulletin & Review. 14*:534–541.

Chinappi, J. (2010). Critical Review of Stanford-Binet Intelligence Scale. Bright hub.

Choudhuri, M. A., Alam K.J., Hasan. R, Rashida , S.A (2005*) Situational Analysis and Assessment of Education for Children with Disabilities in Bangladesh, South Asia, East Asia and South Africa Disability Knowledge and Research Programme.* (CSID) BD.

Church A.T. (2001). Personality measurement in cross-cultural perspective. *Journal of Personality, 69*,979-1006.

Chun, K., Cobb, S., & French, J. R. (1975). *Measures for psychological assessment: A guide to 3,000 original sources and their applications*. Ann Arbor, MI: Survey Research Center, Institute for Social Research.

Clifford, E. (2008). Visual-spatial processing and mathematics achievement: The predictive ability of the visual-spatial measures of the Stanford-Binet intelligence scales, Fifth Edition and the Wechsler Intelligence Scale for Children-Fourth Edition (doctoral dissertation). Retrieved from Proquest Dissertations And Theses.Section 0203, Part 0525 195 pages; [Ph.D.dissertation].United States -- South Dakota: University of South Dakota; Publication Number: AAT 3351188.

Coleman, M. R., Buysse, V., & Nietzel, J. (2006). *Recognition and response: An early intervening system for young children at-risk for learning disabilities. Full report.* Chapel Hill: The University of North Carolina, Frank Porter Graham Child Development Institute.

Collingwood, J. (2008). Intelligence Quotient Rising Worldwide. *Psych Central*. Retrieved on March 8, 2012, from http://psychcentral.com/lib/2008/intelligence-quotient-rising-worldwide/.

Cortina, J.M., (1993). What Is Coefficient Alpha? An Examination of Theory and Applications. *Journal of Applied Psychology, 78*(1), 98-104.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16,* 297-334.

Cronbach, L. J. (1970). *Essentials of Psychological Testing*. New York: Harper and Row.

Daniels, M., Devlin, B., & Roeder, K. (1997). *Of genes and IQ.* NY: Springer.

Daniel Mont (2007).*Measuring Disability Prevalence. Social protection* , Disability & Development Team HDNSP. The World Bank.

Devlin, B., Feinberg, S.E., Resnick, D.P., & Roeder, K., (1997). *Intelligence, genes and success.* NY: Copernicus.

Dickens, W. T., & Flynn, J. R. (2006). Black Americans reduce the racial IQ gap: Evidence from the standardization samples. *Psychological Science, 17,* 913–920.

Domino, G. (2000). *Psychological testing: An introduction*. NJ: Prentice-Hall.

Douglas, N. (2006). *Males have greater g: Sex differences in general mental ability from 100,000 17- to 18-year-olds on the Scholastic Assessment Test.* The University of Western Ontario, Retrieved fromwww.sciencedirect.com.

Duncan, J., Seitz, R.J., Kolodny, J., Bor, D., Herzog, H., Ahmed, A., Newell, F.N., Emslie, H. (2000). A neural basis for general intelligence. *Science*, 289, 457-460.

Embretson, S. E. (1996). The new rules of measurement. *Psychological Assessment, 8,* 341-349.

Embretson, S. E., & Hershberger, S. (1999). *The new rules of measurement: What every psychologist and educator should know.* NJ: Erlbaum.

ESTEEM (2002). *Educating Children in Difficult Circumstances: Children with Disabilities.* Research Series CSID, Directorate of Primary Education, Primary and Mass Education Division.

Emily J. K., & Robert E. L. (1987). *Concurrent validity of the Stanford-Binet fourth edition and K-ABC for head start children.* USA: Southern Illinois University.

Evans, J. J., Floyd. R G., McGraw, K. S., & Leforgee, M. H. (2001). The relations between measures of Cattell-Horn-Carroll (CHC) cognitive abilities and reading achievement during childhood and adolescence. *School Psychology Review 31*(2), 246—262.

Fernandez-Ballesteros, R. (2003).*Encyclopedia of Psychological Assessment*. Edited by London: Sage, (2 vols.) London ; Thousand Oaks, Calif. : SAGE Publications .

Fletcher, J. M., Foorman, B. R., Boudousque, A., Barnes, M. A., Schatschneider, C., & Francis, D.J . (2002). Assessment of reading and learning disabilities: A research-based intervention-oriented approach. *Journal of School Psychology, 40,* 27-63.

Flanagan, D. P (2000). Wechsler-based CHC cross-battery assessment and reading achievement: Strengthening the validity of interpretations drawn from Wechsler test scores. *School Psychology Quarterly, 15*, 295-329.

Flanagan, D. P., Genshaft, J. L., & Harrison, P.L. (Eds.). (1997). *Contemporary intellectual assessment: Theories. tests and issues.* NY: Guilford.

Flynn, J. R. 1994. IQ gains over time. In R. J. Sternberg (ed.). *Encyclopedia of Human Intelligence*. New York: MacMillan, pp. 617-623.

Flynn, James R.(2007) *What is Intelligence ? Beyond the Flynn Effect*: Cambridge University Press.

Ford, L., & Dahinten, V. S. (2005). Use of intelligence tests in the assessment of preschoolers. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment* (pp. 487-503). NY: The Guilford Press.

Francois, G.R. (1995). *Of Children: An introduction to child development* (7th ed.). NY: Wadsworth Publishing Company.

Framingham, J. (2011). What is Psychological Assessment?. *Psych Central*. Retrieved from http://psychcentral.com/lib/2011/what-is-psychological-assessment/

Galton, F. (1884). Measurement of character. *Fortnightly Review, 36,* 179-185.

Gardner, H., Kornhaber, M. L., & Wake, W. K. (1996). *Intelligence: Multiple perspectives.* CA: Wordsworth.

Gardner, H. (1983). Frames of mind: The theory of multiple intelligences. NY: Basic Books.

Geisinger, K.F., Spies, R.A, Carlson, J. F., & Plake, B. S. (Eds.) (2007). *The seventeenth mental measurements yearbook.* Lincoln, NE: Buros Institute of Mental Measurements.

George, D., & Mallery, P. (2003). SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.). Boston: Allyn & Bacon.

Goldman, B. A., & Saunders, J. L. (1995). *Directory of unpublished experimental mental measures* (volume 6). New York: Behavioral Publications.

Gregory, R.J. (2007). *Psychological testing: History, principles, and applications* (5th ed.). Boston, MA:Pearson.

Gustafsson, J. E. (1984,1992). A unifying model for the structure of intellectual abilities. *Intelligence, 8,* 179-203.

Guidelines for cognitive and pilot testing of questions for use in surveys (2010). ESCAP Project On Improving Disability Measurement And Statistics In The Asia, Pacific Region, December, Statistics Division, Economic and Social Commission for Asia Pacific Region.

Hambleton, R. K. (1994).Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment 10*, 229-244.

Hambleton, R. K. (2005). The next generation of the ITC test translation and adaptation guidelines. *European Journal of Psychological Assessment, 17*(3), 164-172.

Hambleton, R. K., Merenda, P., & Spielberger, C. (2005). Adapting educational and psychological tests for cross-cultural assessment. NJ: Lawrence S. Erlbaum Publishers.

Herrnstein, R., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life.* NY: Free Press.

Hersen. *(2004). Comprehensive Handbook of Psychological Assessment*. New York: Wiley.

Hollinger, C.L., & Baldwin, C. (1990). *Comparing scores on the Stanford-Binet, Fourth Edition with the WISC-R for exceptional children*. Ohio: Cleveland State University.

Horn, J.L. (1994). Theory of fluid and crystallized intelligence. In R. J. Stemberg (Ed). *Encyclopedia of human intelligence* (pp. 443-451). NY: Macmillan.

Horn, J. L., & Noll, J. (1997). human cognitive capabilities: Gf-Gc theory. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories; tests and issues* (pp.53-91). NY: Guilford Press.

Horn, J. L. (1985). Remodeling old models of intelligence. In B. B. Wolman , *Handbook of Intelligence. Theories, Measurements, and Application.* (pp. 267-300).NY:Wiley.

Horn, J. L., & Noll, J. (1994). A system for understanding cognitive capabilities: A thoery and the evidence on which it is based. NJ: Ablex.

Huq, S (1996). The validity of five subtests of Stanford Binet intelligence Scale (1986 revision). *Dhaka Univ. J.Psychol.Vol.20*: pp.43-50.

Huq, S (1992). Determination of Reliability and Validity of Stanford Binet intelligence Scale and the Construction of Norm for use in Bangladesh Unpublished Doctoral Thesis , Department of Psychology, University of Dhaka, Dhaka.

Huq, S (1989). Item analysis of the five subtests of Stanford Binet intelligence Scale (translated version in Bengali) for use in Bangladesh . *AFMR:* Proceedings of 9th Asian conference on mental retardation. Thailand Bangkok .

Huq, S (1980).Standardization of Wechsler Intelligence Scale for Children (Translation and Adaptation in Bengali) for use in Dhaka city. *Bangladesh J. Psychol. Vol.6.* University of Dhaka, Department of Psychology.

Hurlow , S. C.(2011). *Item Fairness of the Nonverbal Subtests of the Stanford-Binet Intelligence Test, Fifth Edition, in a Latina/o Sample*. George Fox University, 47 pages; 3485301.

Individuals with Disabilities Education Act of 1997, 20 U.S.C 1400 et seq: U.S [CL1]. Statutes at Large, 104, 1103—1151, re authorized 1997.

International Test Commission. (2001). International guidelines for test: Version 2000. London, England: Author. Retrieved from http://www.intestcom.org/itc-projects.htm#IT C%2OGuidelines%20on%20Test%20Use

Jean-Marc,W.(2008). APPS Assistant Director: Center for Assessment & Testing. (jwise@campus.fsu.edu, 644-3541).

Johnson, J. A., & D'Amato, R. C. (2006). Review of the Stanford-Binet Intelligence Scales: Fifth Edition. In R. S. Spies & B. S. Plake (Eds.), *The sixteenth mental measurements yearbook,* (pp. 976–979). NE: Buros Institute of Mental Measurements.

Kamphaus R. W.(1993). *Clinical assessment of children's intelligence*. Boston: Allyn & Bacon.

Kaplan, R. M., & Sacuzzo, D. P. (2001). *Psychological testing: Principles, applications and issues* (5th Ed.). CA: Wadsworth/Thomson.

Kaufman, A. S., & Kaufman, N. L. (1993). *Administration and scoring manual for the Kaufman brief intelligence test (K-BIT).* MN: American Guidance Service.

Keeffe , J., Baker., S, Booth., J., Goujon., N, Edmonds., T, Huq N. L, & Quaiyum, A. (2011)Rapid Assessment of Disability News letter, Issue 3, February, ICCDRB, Dhaka.

Kline, P (1999, 1996). *The Handbook of Psychological Testing*. N Y, USA: Routledge.

Kline. R. B., Graham, S. A., & Lachar, D. (1991) Are IQ scores valid for children who are poor readers? *Psychological Assessment, 5.* 400—407.

Kolar, G.M., (2001). *A Literature Review and Critical Analysis of the Concurrent Validity of the Differential Ability Scales and the Cognitive Assessment System.* Menomonie: University of Wisconsin –Stout.

Kulshrestha, S. K. (1971). Stanford Binet Intelligence Scale and manual. Hindi Adaptation (Under third revision) *Manas Seva Sansthan Prakashan*, Allahabad.

Kush, J. C. (2005). Review of the Stanford-Binet Intelligence Scales: Fifth Edition. In R. A. Spies & B. S. Plake (Eds.), *The sixteenth mental measurements yearbook* (pp. 979–984). Lincoln: University of Nebraska.

Laurent, J., Swerdlik, M., & Ryburn, M. (1992). Review of validity research on the Stanford-Binet Intelligence Scale-Fourth Edition. *Psychological Assessment, 4*, 102-112.

Lawler, J.M. (1978). *Reliability and validity of IQ tests*. NY: International Publishers.

Lee, D., Reynolds, C. R., Willson, V. L. (2003). Standardized test administration: why bother? *Journal of Forensic Neuropsychology, 3*, 55–81.

Lincoln, E. A. (2010). The reliability of the Stanford-Binet scale and the constancy of intelligence quotients. *Journal of Educational Psychology. 18(*9), 621-626.

Logsdon, A. (2011). *Learning Styles - Learn about the Verbal Linguistic Learning Style.* Learning Styles. About.com Guide .

Maddox, T (2003).*Tests : A Comprehensive Reference for Assessments in Psychology, Education and Business* / 5th ed. Kansas City : Test Corp. of America. 2003.

Madsen. (1934). The Reliability and Validity of the Stanford-Binet Tests when Administered by Student Examiners. *Journal of Educational Research, 28*(4), page no. 461.

Matarazzo, J. D. (1990). Psychological assessment versus psychological testing: Validation from Binet to the school, clinic, and courtroom. *American Psychologist, 45*, 999-1017.

Matthews, L. (2003). *Best practices and technical issues in cross-lingual, cross-cultural assessments: An evaluation of a test adaptation.* USA Ohio University

McArdle, J. J. & Woodcock, R. W. (1997). Expanding test-retest designs to include developmental time lag components. *Psychological Methods, 2*(4) 403-435.

McCauley (2011) Intelligence, race, and genetics. *American Psychologist, 60*, 46–59.

McGrew, K. S., & Hessler, G. L. (1995). The relationship between the WJ-R Gf-Gc cognitive clusters and mathematics achievement across the life-span. *Journal of Psych-educational Assessment, 13*, 21-38.

McGrew, K S., Keith, T. Z., Flanagan, D. P., & Vanderwood, M. (1997). Beyond g: The impact of Gf Gc specific cognitive abilities research on the future use and interpretation of intelligence tests in the schools. *School Psychology Review, 26,* 189-210.

McGrew, K. S., & Woodcock, R. W. (2001). *Technical manual.* IL: Riverside Publishing.

Mc Grew, K. (2010). The Wechsler-like IQ subtest scaled score metric: The potential for misuse, misinterpretation and impact on critical life decisions. *Institute for Applied Psychometrics (IAP)* January 05.

McNemar (1942). *The revision of the Stanford–Binet Scale*. Boston: Houghton Mifflin.

Meloff, R (1987). *The Stanford-Binet Intelligence Scale, fourth edition: a Descriptive study.* University of Alberta : Department of Educational Psychology Publisher.

Mellenbergh, G.J. (2008). Chapter 10: Surveys. In H.J. Adèr & G.J. Mellenbergh (Eds.) (with contributions by D.J. Hand), *Advising on Research Methods: A consultant's companion* (pp. 183-209). The Netherlands: Johannes van Kessel Publishing.

Miller, D. C. (1991). Handbook of research design and social measurement (5th ed.). New Park, CA: Sage.

Mitra, N. K., Nagaraja, H. S., Ponnudurai, G., & Judson, J. P. (2009). The Levels Of Difficulty And Discrimination Indices In Type A Multiple Choice Questions Of Pre-clinical Semester 1 Multidisciplinary Summative Tests IeJSME 2009: 3 (1): 2-7

Department of Human Biology, International Medical University, Bukit Jalil, 57000 Kuala Lumpur, Malaysia.

Munoz-Sandoval, Cummins, Alvarado, and Ruef . Assessing Lexical Development in Bilingual Babies and Toddlers. *International Journal of Bilingualism November 1998 vol. 2 no. 3 347-37.* Sage Journals. doi: 10.1177/13670069800200305.

National Education Policy 2010, Bangladesh, Ministry of Education Chapter 18. Special Education, Health and Physical Education, Scout, Girls' Guide and Bratachari. p.43 2010/06/01/

Neisser, U. (1998). *The rising curve: long-term gains in IQ and related measures.* Washington, D.C: American Psychological Association.

Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., & Ceci, S. J.(1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*, 77-101.

No Child Left Behind Act (2001), Pub.L. No. 107–110, 115 Stat. 1435 (2002).

Overton, T. (1992). *Assessment in special education an applied approach.* NY: Maxwell Macmillan Company.

Peterson, J. (1925). *Early conceptions and tests of intelligence*. NY: World Book.

Phelps, R.P. (2008). Educational achievement testing fallacies. In R.P. Phelps (Ed.), *Correcting fallacies about educational and psychological testing.* Washington, DC: American Psychological Association.

Pintner, R., & Paterson, D.G. (1917). *A scale of performance tests*. NY: Appleton.

Popham, J. (1999). Why standardized tests don't measure educational quality. *Educational Leadership, 56*(6), 8-15.

Pomplun, M. (2006). *The Construct Validity of the Stanford-Binet 5 Measures of Working Memory.* Michael Custer : Riverside Publishing.

Rasch, G. (1960, 1980). *Probabilistic models for some intelligence and attainment tests.* (Copenhagen, Danish Institute for Educational Research), expanded edition (1980) with foreword and afterword by B.D. Wright. Chicago: The University of Chicago Press.

Reckase, M. D. (1989). Adaptive testing: The evolution of a good idea. *Educational Measurement: Issues & Practice, 8,* 11–15

Reckase, M. D. (1996). Test construction in the 1990s: Recent approaches every psychologist should know. *Psychological Assessment, 8,* 354-359.

Reid,D.K., Hresko, W. P., & Swanson, H. L. (1996). *Cognitive approaches to learning disabilities*. TX: Pro-Ed.

Reschly, D., Tilly, D., & Grimes, J. (1999). *Special education in transition.* Sopris West:Longwood, CO.

Reynolds, C. R. (1985). Critical measurement issues in learning disabilities. *Journal of Special Education, 18,* 451-476.

Reynolds, C. R., Livingston, R. A., & Willson, V. L. (2006). *Measurement and assessment in the classroom.* Boston: Allyn and Bacon.

Richardson, K. (2002). What IQ tests test. Open university : SAGE journal.

Robinson, N. M. (2002). *Assessing and advocating for gifted students: Perspectives for school and clinical psychologists* (RM02166). Storrs, CT: The National Research Center on the Gifted and Talented, University of Connecticut.

Roberts, C. & Cox, M. (2005). Medieval ancestors measured up to our height standards. British Archaeology, 84:51.

Rodriguez, C, Treacy, L & Sowerby,P & Murphy, L (1998).Applicability of Australian Adaptations of Intelligence Test in Newzealand with a Dunedin Sample of Children. Newzealand Journal of Psychology.27 (1).pp. 4-13.

Roid, G. H. (2003). *Stanford-Binet Intelligence Scales.* IL: Riverside Publishing.

Roid, G. H. (2003a). *Stanford-Binet Intelligence Scales, Fifth Edition.* Itasca, IL: Riverside Publishing.

Roid, G. H. (2003b). *Stanford-Binet Intelligence Scales, Fifth Edition, examiner's manual.* Itasca, IL:Riverside Publishing.

Roid, G. H. (2003c). *Stanford-Binet Intelligence Scales, Fifth Edition, interpretive manual: Expanded guide to the interpretation of SB5 test results.* Itasca, IL: Riverside Publishing.

Roid, G. H. (2003d). *Stanford-Binet Intelligence Scales, Fifth Edition, technical manual.* Itasca, IL: Riverside Publishing.

Roid, G. H., & Miller, L. J. (1997). *Leiter International Performance Scale-Revised.* IL: Stoelting.

Rousson, V., Gasser, T., & Seifer, B. (2002). "Assessing intrarater, interrater and test–retest reliability of continuous measurements" *Statistics in Medicine*, 21(22): 3431–3446.

Sattler, J. M. (1992). *Assessment of children: Behavioral and clinical applications* (3rd ed.). , San Diego: Publisher, Inc.

Sattler, J. M. (2001). *Assessment of children cognitive applications* (4th ed.) San Antonio: Sattler.

Sattler, J. M. (l988). *Assessment of children* (3rd ed.). CA: Author.

Snow, R. E., Kyllonen, P. C., & Marshalek, B. (1984). The topology of ability and learning correlations. In R. J. Sternberg (ed.), *Advances in the psychology of human intelligence* (pp. 47-103). NJ: Erlbaum.

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

Spearman, C. (1927). *The Abilities of Man.* London: MacMillan.

Standards for Educational and Psychological Testing. Committee to Develop Standards of Educational Psychological Testing. Washington, D.C.: *American Psychological Association.*

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of intelligence.* Cambridge: Cambridge University Press.

Sternberg, R. J. (2000). *Handbook of intelligence*. NY: Cambridge University Press.

Sternberg, R. J., & Grigorenko, E. L. (2005). Cultural explorations of the nature of intelligence. In Alice F. Healy (Ed.), *Experimental cognitive psychology and its applications* (pp. 225–235). Washington, DC: American Psychological Association.

Sternberg, R.J, & Kaufman, J.C. (1998). Human abilities. *Annual Review of Psychology, 49*, 479–502.

Terman, L. M. (1916). *The measurement of intelligence: An explanation of and a complete guide for the use of the Stanford revision and extension of the Binet-Simon Intelligence Scale.* Boston: Houghton Mifflin.

Terman, L. M., & Merrill, M. A. (1937). *Measuring intelligence*. Boston: Houghton Mifflin.

Terman, L. M., & Merrill, M. A. (1960). *Stanford-Binet Intelligence Scale: Manual for the third revision Form L-M.* Boston: Houghton Mifflin.

Terman, L. M., & Merrill, M. A. (1973). *Stanford-Binet Intelligence Scale: Manual for the third revision Form L-M* Boston: Houghton Mifflin.

Thompson, B. (1984). *Canonical correlation analysis: Uses and interpretation.* Newbury Park: Sage.

Thompson, D. L. (2003). Review of the Multidimensional Aptitude Battery, Second Edition. In B. Plake, J. Impara, & R. Spies (Eds.), The *fifteenth mental measurements yearbook* (pp. 603–607). Lincoln: University of Nebraska.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-Binet Intelligence Scale* (4th ed.). IL: Riverside.

Thorndike, R. M. (1990). *A century of ability testing.* Chicago : Riverside.

Thorndike, R. M. and D. F. Lohman (1994). *A Century of Ability Testing*. Chicago, Riverside Publishing Company.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986a). Guide for Administering and Scoring - Stanford Binet Intelligence Scale (4th ed.). IL: Riverside Publishing.

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

Tucker, S. K. (1991). A validation study of the General Purpose Abbreviated Battery of the Stanford-Binet-Fourth Edition used in the reevaluation of learning-disabled students. Dissertation Abstracts International, 51(8.B), 4098—4099.

Turner, S. M., DeMers, S. T., Fox, H. R., & Reed, G.M. (2001). APA's guidelines for test user qualifications. *American Psychologist, 56,* 1099-1113.

United Nations Development Programme(UNDP) Report (2001). Wikipedia.

U.S. Census Bureau, (2001). *Census 2000 Summary File 1 United States.* Washington, DC: U.S. Census Bureau.

U.S. Department of Education (2004, August 26). IDEA '97: *The Individuals with Disabilities Education Act Amendments of 1997.* Retrieved from http://www.ed.gov/offices/OSERS/Policy/IDEA/index.html /

U.S. Department of Education. (2004, September 16). *Stronger accountability: Testing for results: Helping families, schools, and communities understand and improve student achievement*. Retrieved February 15, 2006, from http://www.ed.gov/nclb/accountability/ayp/testingforresults.html

Vijer, F.V., & Hambleton, R. K. (1996). Translating tests: some practical guidelines. *European Psychologist, 1,* 89-99.

Vincent, E., Glaub, R., & Kamphaus, W. (1999, 2011). Construction of a Nonverbal Adaptation of the Stanford-Binet Fourth Edition. *Educational and Psychological Measurement 51*(1), 231-241.

Vernon, P E.(1961) . *The Structure of Human Abilities.* London: Methuen.

Wagner, R. K. (2000). Practical Intelligence. In A. E. Kazdin (Ed). *Encyclopedia of psychology: Vol. 6* (pp. 266–270). Washington, DC: Oxford University Press.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children-Third Edition manual.* TX: Psychological Corporation.

Williams, P. (1996). *The relationship between arousal, affect, and cognition.* Unpublished doctoral dissertation, Auburn University, AL.

Woodcock, R. W. (1990). Theoretical foundations of the WJ-R measures of cognitive ability. *Journal of Psycho educational Assessment, 8,* 231-258.

Woodcock, R. W., & McGrew, K, S. (1997). *Woodcock- Johnson III Tests of Achievement.* IL: Riverside Publishing.

Woodcock, R. W., & Munoz-Sandoval, A. F. (2001). *Comprehensive manual-Woodcock-Munoz Language Survey Normative Update.* IL: Riverside Publishing. Wolf, T.H. (1973). *Alfred Binet.* Chicago: University of Chicago Press.

Wolf, T. H. (1973).Alfred Binet. Chicago: University of Chicago Press.

World Health Organization (WHO), 2012. Wikipedia.

Wright., & Uneacre. (1999). Fundamental measurement in psychology. In S.E. Embretson & S.L. Hershberger (Eds.), *The new rules of measurement: What every psychologist and educator should know* (pp. 65-104). NJ: Lawrence Erlbaum Associates.

Appendix

Appendix 1: Letter from Supervisor

শিক্ষা ও গবেষণা ইন্‌স্টিটিউট
ঢাকা বিশ্ববিদ্যালয়
ঢাকা–১০০০, বাংলাদেশ
ফোন: ৯৬৬১৯২০-৭৩/
ই-মেইল:

Institute of Education & Research (IER)
University of Dhaka
Dhaka-1000, Bangladesh
Fax : 880-2-8615583
E-mail :

_____
_____
_____
_____
_____

Subject: Permission to administer the Stanford - Binet Intelligence Scale on the students.

Dear Sir/Madam

This is to inform you that Nigar Sultana, is undergoing her Ph.D research from the Department of Special Education, Institute of Education and Research, University of Dhaka.

The title of her research study is "**Standardization of Stanford –Binet Intelligence Scale (fifth edition) for use in Bangladesh".**

On behalf of the above context Nigar Sultana needs to administer the scale on at least 20 to 25 students from ages six to sixteen years in different schools.

The documents will be kept confidential.

The department will be benefited from your kind co-operation.

Thanking you

Yours sincerely

Sharmin Huq

(Dr.Sharmin Huq)
Supervisor

DR. SHARMIN HUQ
Associate Professor
Institute of Education & Research
University of Dhaka

Appendix 2: Letter from DEO

গনপ্রজাতন্ত্রী বাংলাদেশ সরকার
জেলা শিক্ষা অফিসারের কার্যালয়
চট্টগ্রাম ।

স্মারক নং . ১৩১৮/৫/০৭ তারিখ১৫/০৫/২০০৭

প্রেরকঃ জেলা শিক্ষা অফিসার ,
চট্টগ্রাম ।

প্রাপকঃ অধ্যক্ষ / প্রধান শিক্ষক / শিক্ষিকা ,
..... উচ্চবিদ্যালয় / কলেজ
ডাকঘর..... থানা/ উপজেলা....., জেলা- চট্টগ্রাম ।

বিষয়ঃ গবেষণা কার্যে সহায়তা করণ প্রসংগে।

সূত্র ঃ শিক্ষা গবেষণা ইনষ্টিটিউট ঢাকা এর অনুরোধ পত্র।

উপর্যুক্ত বিষয় ও সুত্রের আলোকে জানান যাচ্ছে যে, তাঁর বিদ্যালয়ে জনাব নিগার সুলতানা এর পিএইস ডি কার্যে (সরকারী)তথ্য সংগ্রহে সার্বিক সহযোগীতা প্রদানে অনুরোধ জানানো হইল ।

(হোসনে আরা বেগম)
জেলা শিক্ষা অফিসার,
চট্টগ্রাম ।

স্মারক নং ১৩১৮/ ৫/ ০৭ তারিখ ঃ ১৫/০৭/২০০৭

সদয় অবগতি ও প্রয়োজনীয় কার্যার্থে অনুলিপি প্রেরণ করা হলো ঃ

১ । মহাপরিচালক, মাধ্যমিক ও উচ্চ শিক্ষা অধিদপ্তর, বাংলাদেশ, ঢাকা ।
২ । ড: শারমীন হক ,সহযোগী অধ্যাপক শিক্ষা গবেষণা ইনষ্টিটিউট ঢাকা
৩ । জনাব নিগার সুলতানা প্রভাষক শিক্ষা গবেষণা ইনষ্টিটিউট ঢাকা
৪ । উপ-পরিচালক,মাধ্যমিক ও উচ্চ শিক্ষা পরিদপ্তর, চট্টগ্রাম অঞ্চল, চট্টগ্রাম ।
৫ । সংরক্ষন নথি ।

১৫/৭/০৭

(হোসনে আরা বেগম)
জেলা শিক্ষা অফিসার,
চট্টগ্রাম ।

Appendix 3: Letter of Seminars

Seminar 1

INSTITUTE OF EDUCATION AND RESEARCH
Dhaka University

Seminar on Ph. D. Thesis

Ms. Nigar Sultana, Lecturer, Department of Special Education and Ph. D. fellow of IER will give her first seminar on the proposed thesis titled, "Standardization of Stanford-Binet Intelligence Scale: Fifth Edition (Translation and Adaptation for use in Bangladesh)". The seminar will be held on 12th February 2008 at Room No. 209, IER at 12:30 pm.

All the faculty members along with the research students are cordially invited to attend the seminar.

Date: 6.2.08

M. Nazmul Haq
Chairman
Seminar and Training Committee

Seminar 2

Ph.D Seminar

Date: 08.03.2011

Dear Colleagues,

I would like to inform you that Ms. Nigar Sultana, Assistant Professor and Chairman, Department of Special Education, IER, University of Dhaka will present the progress of her Ph.D research work on 15th March 2011 at 10:00 am. in Room No. 209 (Conference room). The title of her study is "Standardization of Stanford Binet Intelligence Scale (Fifth Edition, 2003) for use in Urban Bangladesh".

You are requested to attend the seminar.

Thanking you

Professor Dr. Delwar Hossain
Chairman
Seminar Committee
IER, University of Dhaka

Appendix 4: Name of the Experts for Content Validity

Dr. Sultana Sarwat Ara Zaman Professor Emeritus Department of special education Institute of Education and Research University of Dhaka	**K.S.M. Ilyas** Professor Department of Psychology University of Dhaka
Rokeya Begum Professor Department of Educational Evaluation and Research Institute of Education and Research University of Dhaka	**Dr. Shaheen Islam** Professor and Chairman Department of Educational Psychology & Counseling University of Dhaka
Dr Shirin Zaman Munir Expert in Special and Disability Issues Part - Time Faculty Member Department of Psychology University of Dhaka	**Dr. A.K.M. Rezaul Karim** Associate Professor Department of Psychology University of Dhaka

Appendix 5: Declaration of Seventh Division – Rangpur

The Daily Star

Tuesday, January 26, 2010

Rangpur made 7th division: Staff Correspondent

The government has declared Rangpur the nation's seventh division. The people of Rangpur have long been demanding a separate division in order to encourage development in its impoverished areas. The decision was announced yesterday by the National Implementation Committee for Administrative Reorganization (Nicar), which is chaired by Prime Minister Sheikh Hasina. The Prime Minister's Press Secretary Abul Kalam Azad said that creating the new division fulfills an election pledge of the Awami League. The new division comprises eight districts that were formerly under Rajshahi division -- Rangpur, Kurigram, Dinajpur, Lalmonirhat, Gaibandha, Nilphamari, Thakurgaon and Panchagarh. The total population of Rangpur division is 1, 4847,000.The press secretary said that the Cabinet decided to create the new division during a meeting on July 13 last year. A committee was subsequently formed and its report on creating a new division was submitted to the government on July 21. Nicar has also decided to establish four police stations in the Barisal Metropolitan area. New stations will be located in Kaunia, Airport, Bandar and Babuganj, and Kotwali police station will be reopened. Bangladesh's six other administrative divisions are Dhaka, Chittagong, Rajshahi, Khulna, Barisal and Sylhet.

News Source: **The Daily Star**

Appendix 6: Distributions of students according to sex, age and demographic variables

Table

Distribution of students showing their sex according to division

Division	Students(M)	Students(F)	Total
Dhaka	275	275	550
Chittagong	275	275	550
Sylhet	275	275	550
Barisal	275	275	550
Khulna	275	275	550
Rajshahi	275	275	550
Total	1650	1650	3330

Table

Distribution of students showing their sex according to Districts, Thanas, and Metropolitan schools of Bangladesh

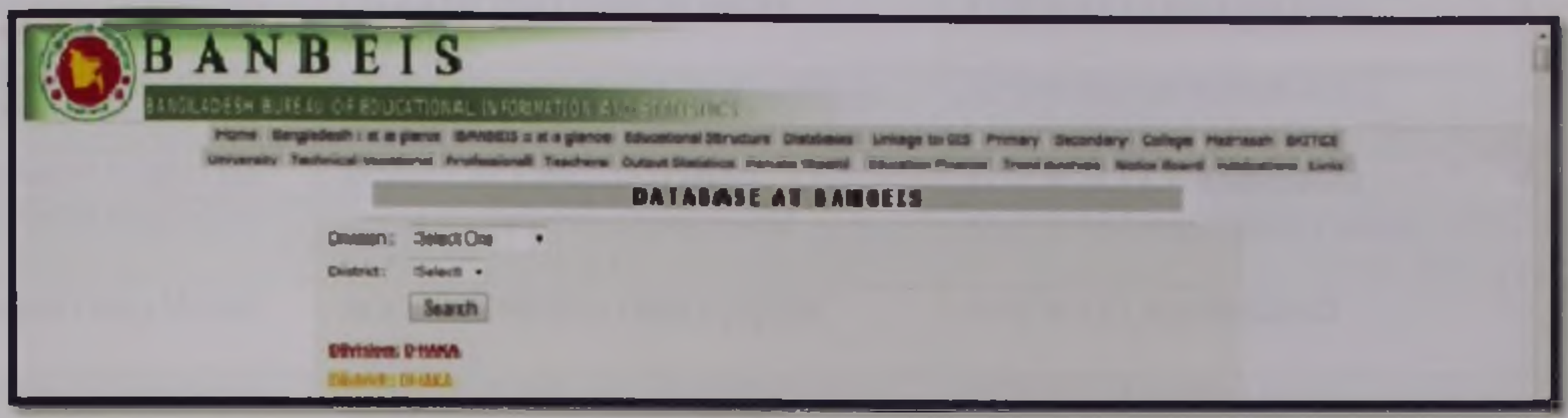
Division	Metropolitan cities	Thanas	Schools (Metropolitan)	Students	
				M	F
Dhaka	Dhaka	7+14=21	30	275	275
Chittagong	Chittagong	11	25	275	275
Barisal	Barisal	5	23	275	275
Khulna	Khulna	1(sadar)	22	275	275
Rajshahi	Rajshahi	1(sadar)	25	275	275
Sylhet	Sylhet	1(sadar)	25	275	275
6	6	40	150	1650	1650
Total				3300	

Table

Distribution of students showing their age, sex according to Divisions of Bangladesh

Age Group	Divisions												Total
	Barisal		Chittagong		Dhaka		Khulna		Rajshahi		Sylhet		
	M	F	M	F	M	F	M	F	M	F	M	F	
6	25	25	25	25	25	25	25	25	25	25	25	25	300
7	25	25	25	25	25	25	25	25	25	25	25	25	300
8	25	25	25	25	25	25	25	25	25	25	25	25	300
9	25	25	25	25	25	25	25	25	25	25	25	25	300
10	25	25	25	25	25	25	25	25	25	25	25	25	300
11	25	25	25	25	25	25	25	25	25	25	25	25	300
12	25	25	25	25	25	25	25	25	25	25	25	25	300
13	25	25	25	25	25	25	25	25	25	25	25	25	300
14	25	25	25	25	25	25	25	25	25	25	25	25	300
15	25	25	25	25	25	25	25	25	25	25	25	25	300
16	25	25	25	25	25	25	25	25	25	25	25	25	300
11 groups	**275**	**275**	**275**	**275**	**275**	**275**	**275**	**275**	**275**	**275**	**275**	**275**	**3300**

Appendix 7 : List of Schools of Six Divisions



Division: BARISAL
District: BARISAL

Institute ID	Institute Name	Thana ,Post Office
010651171383001	TALUKDER HAT SCHOOL & COLLEGE	KOTWALI,TALUKDERHAT
010651173391201	PRAFULLA NAHA GIRLS HIGH SCHOOL	KOTWALI,NAIMITRA
010651173391303	DURGAPUR HIGH SCHOOL	KOTWALI,DAKSHINDURGAPUR
010651176171302	CHANDPURA UNION HIGH SCHOOL	KOTWALI,NAIMITRA
010651179551304	HIZAL TALA MOULAVIRHAT HIGH SCHOOL	KOTWALI,HIZAL TALA
010651251851301	CHANDRA MOHAN R.M. HIGH SCHOOL	KOTWALI,CHANDRA MOHAN
010651342851303	CHARBARIA HIGH SCHOOL	KOTWALI,CHARBARIA
010651342931304	LAMCHARI HIGH SCHOOL	KOTWALI,CHARBARIA
010651346161201	KAGASURA SECONDARY SCHOOL	KOTWALI,CHARBARIA
010651346161307	ALHAZ A. MAZID KHAN SECONDARY SCHOOL	KOTWALI,CHAND MOHON
010651346941306	MAHABAZ SECONDARY SCHOOL	KOTWALI,CHARBARIA
010651347791302	PURAN PARA HIGH SCHOOL	KOTWALI,BARISAL
010651349791305	AHABAZ HIGH SCHOOL	KOTWALI,CHARBARIA
010651431691201	SHAHEED ZIAUR RAHMAN SECONDARY SCHOOL	KOTWALI,CHAR BADNA
010651432151307	CHARKARANJI HIGH SCHOOL	KOTWALI,CHAR KARANJI
010651432231206	CHAR KAWAYANAYANI JUNIOR H/S	KOTWALI,CHAR KOWA
010651433311203	PANCHAGRAM SHAMMALANI JR.GIRL'S SCHOOL	KOTWALI),CHAR KOWA
010651433311301	ABDUR RASHID KHAN HIGH SCHOOL	KOTWALI,KHANPUR
010651435473001	KARNOKATI G.R. SCHOOL & COLLEGE	KOTWALI,KARNAKATI
010651435781302	TOFAEL AHMAD HIGH SCHOOL	KOTWALI,CHARKOWA
010651512081304	CHAR HOGLA HIGH SCHOOL	KOTWALI,CHAR HOGLA

Institute ID	Institute Name	Thana ,Post Office
010651512501207	CHATUA CHAR GOPALPUR JUNIOR HIGH SCHOOL	KOTWALI,BUKHAINAGAR
010651513001303	CHARMONAI UNION HIGH SCHOOL	KOTWALI,CHARMONAI
010651513051301	BUKHAINAGER SECONDARY SCHOOL	KOTWALI,BUKHAINAAR
010651518021302	JABED ALI INSTITUTION	KOTWALI,CHARMONAI
010651518021305	RAJAR CHAR SECONDARY SCHOOL	KOTWALI,SAHEBER CHAR
010651601311303	SOLAMONESSA GIRLS HIGH SCHOOL	KOTWALI,AMIRABAD
010651601311304	MOHAMMAD ALI SECONDARY SCHOOL	KOTWALI,AMIRABAD
010651604391206	A. AZIZ KHAN JUNIOR SCHOOL	KOTWALI,ZAGUAHAT
010651604621301	HOGLA MULTILATERAL HIGH SCHOOL	KOTWALI,HOGLA
010651608561302	RUPATALI ZAGUA SECONDARY SCHOOL	KOTWALI,JAGUAHAT
010651608561305	A. WAHED GIRLS HIGH SCHOOL	BARISAL SADAR (KOTWALI), JAGUAHAT
010651694011205	KASHIPUR JUNIOR GIRLS SCHOOL	KOTWALI,GANAPARA
010651694703002	KASHIPUR HIGH SCHOOL & COLLEGE	KOTWALI,KASHIPUR
010651695551202	ELEM UDDIN SHARIF JUNIOR HIGH SCHOOL	KOTWALI,KASHIPUR
010651695551307	CHAHUTPUR SECONDARY SCHOOL	KOTWALI,KASHIPUR
010651696471301	P.R.C. INSTITUTION LAKHUTIA	KOTWALI,LAKHUTIA
010651774851302	NAYAGAON HIGH SCHOOL	KOTWALI,RUIYA
010651779101304	SHOLNA HIGH SCHOOL	KOTWALI,SHOLNA
010651779941303	KARAPUR POPULAR HIGH SCHOOL	KOTWALI,KARAPUR
10651860071302	SHAESTABAD M.H. HIGH SCHOOL	KOTWALI,SHAESTABAD
010651860151301	CHAPILADI GIRLS HIGH SCHOOL	KOTWALI,BUKHAINAGAR
010651861611204	CHAR AICHA JUNIOR SCHOOL	KOTWALI,SHAESTABAD
010651861611208	ENGINEER SYED ABDUL HANNAN JUNIOR SECONDARY SCHOOL	KOTWALI,SAYESTABAD
010651861611305	SHAISTABAD GIRLS HIGH SCHOOL	KOTWALI,SHAESTABAD
010651864511303	SHAFI UDDIN SECONDARY SCHOOL	KOTWALI,CHURAMON
010651868201206	RAMKATY JUNIOR SCHOOL	KOTWALI,CHURAMAN
010651941231305	B.K. HIGH SCHOOL	KOTWALI,SAHEBER HAT
010651943701203	DHOPAKATI JUNIOR HIGH SCHOOL	KOTWALI,PATANG
010651947011201	MOLLRHOWLA JUNIOR GIRLS SCHOOL	KOTWALI,SINGHER KATHI
010651947551304	PATANG HIGH SCHOOL	KOTWALI,PATANG
010651949021302	SHINGHER KATI M.L. HIGH SCH.	KOTWALI,SINGHER KATHI

Division: CHITTAGONG
District: CHITTAGONG

Institute ID	Institute Name	Thana ,Post Office
021510172211301	JULEKHA AMINUR RAHMAN C.C. HIGH SCHOOL	BAKALIA,BAKLIA
021510181101201	MERN SUN JUNIOR HIGH SCHOOL	BAKALIA,BAKLIA
021510181101202	ALHAJ TAJUL ISLAM JUNIOR HIGH SCHOOL	BAKALIA,CHANDAGOAN
021519270861203	SOUTH-WEST BAKALIA JUNIOR HIGH SCHOOL	CHANDGAON,CHAWKBAZAR
021519270861204	SANOWARA ISLAM BOYS JONIOR SCHOOL	CHANDGAON,CHANDGAON
021519270861301	CHANDGAON N. M. C. MODEL HIGH SCHOOL	CHANDGAON,CHANDGAON
021519270861302	NURUL ISLAM MUNICIPAL GIRLS' HIGH SCHOOL	CHANDGAON,CHANDGAON
021519299081301	BAKALIA HIGH SCHOOL	CHANDGAON,CHAK BAZAR
021519312211301	SONOWARA HIGH SCHOOL	CHANDGAON,CHANDGAON
021519312211302	BAKALIA MODEL GIRLS' HIGH SCHOOL	CHANDGAON,CHAK BAZAR
021519312211303	BAKALIA GOVT.LABORATORY HIGH SCHOOL	CHANDGAON,CHAK BAZAR
021519323311302	DAKKHIN PASCHIM BAKALIA HIGH SCHOOL	BAKALIA,CHAK BAZAR
021519323321203	HASNE HENA JUNIOR GIRLS SCHOOL	CHANDGAON,CHAW BAZAR
021519323321301	CHAR CHAKTAI CITY CORP. HIGH SCHOOL	CHANDGAON,SADAR-4000
021519323481301	MD. KAMAL UDDIN HIGH SCHOOL	CHANDGAON,CHANDAGOAN
021520373001303	MARINE ACADEMY HIGH SCHOOL	BANDAR,MARINE ACADEMY
021520373201304	PORT AUTHORITY HIGH SCHOOL	BANDAR,PORT AUTHORITY
021520375101302	NIMTALA HIGH SCHOOL	BANDAR,BANDAR
021520385371301	HALISHAHAR MEHER AFZAL HIGH SCHOOL	BANDAR,BANDAR
021520385371302	HALISHAHAR AHMED MEAH CITY CORP. G.H.SCH	BANDAR,BANDAR
021520385371303	HALISHAHAR BEGUMJAN HIGH SCHOOL	BANDAR,BANDAR
021520385371304	CHITTAGONG PORT AUTHORTY GIRLS HIGH SCHOOL	BANDAR,BANDAR
021520391611202	NAVY HIGH SCHOOL	BANDAR,SAILS COLONY
021520391611301	SOUTH HALISHAHAR HIGH SCHOOL	BANDAR,SAILS COLONY
021520391613003	BEZPA PUBLIK SCHOOL & COLLEGE	BANDAR,BANDAR
021520408871301	EASTERN REFINERY MODEL SCH.	BANDAR,PATENGA

Institute ID	Institute Name	Thana ,Post Office
021520408871304	CHITTAGONG STILLS MILLS HIGH SCHOOL	BANDAR,UTTAR PATENGA
021528152571301	CHITTAGONG POLICE INSTITUTION	DOUBLE MOORING,DAMPARA
021528181101302	ISPAHANI PUBLIC SCHOOL	DOUBLE MOORING,G.P.O
021528193311312	CHITTAGONG COLLEGIAT SCHOOL	DOUBLE OORING,Z.P.O.CTG
021528193601302	RAILWAY EMPLOYEES GIRLS HIGH SCHOOL	DOUBLE MOORING,CTG.GPO
021528197731301	HATEY KHARI HIGH SCHOOL	DOUBLE MOORING,BANDAR
021528197731307	KALAKAKOLI HIGH SCHOOL	DOUBLE MOORING,BANDAR
021528197731308	BANGLADESH BANK COLONY HIGH SCHOOL	DOUBLE MOORING,BANDAR
021528197731310	AGRABAD BALIKA BIDAYALAY	DOUBLE MOORING,BANDAR
021528204131303	T & T HIGH SCHOOL	DOUBLE MOORING,BANDAR
021528204131304	AGRABAD GOVT. COLONY HIGH SCHOOL	DOUBLE MOORING,BANDAR
021528242001304	SILVER BELLS GIRLS HIGH SCHOOL	DOUBLE MOORING,BANDAR
021541011491302	BANGLADESH MAHILA SAMITY GIRL'S HIGH SCHOOL	KOTWALI,GPO
021541011491304	BAGMONIRAM A.R CITY CORP. BOY'S HIGH SCHOOL	KOTWALI,DAMPARA
021541019201301	RAILWAY PUBLIC SCHOOL	KOTWALI,CHITTAGONG SADAR
021541024223003	KAPASGOLA CITY CORP. GIRLS HIGH SCHOOL & COLLEGE	KOTWALI,CHAKBAZAR
021541024971302	GULEZAR BEGUM MUSLIM GIRL'S HIGH SCHOOL	KOTWALI,CHAKBAZAR
021541025471301	KAZEM ALI HIGH SCHOOL	KOTWALI,CHITTAGONG
021541025471304	GOVT. HAJI MD. MOHSIN HIGH SCHOOL	KOTWALI,CHAK BAZAR
021541025471305	CHITTAGONG GOVT. HIGH SCHOOL	KOTWALI,CHAK BAZAR
021541036301302	KADAM MOBARAK CITY CORP. HIGH SCHOOL	KOTWALI,ANDARKILLA
021541036301303	JAMALKHAN K.K. CITY CORPORATION GIRL'S HIGH SCHOOL	KOTWALI,ANDARKILLA
021541036961301	SALEH JAHUR HIGH SCHOOL	KOTWALI,CHITTAGONG SADAR
021541036961305	DR. KHASTAGIR'S GOVT. GIRLS HIGH SCHOOL	KOTWALI,SADAR
021541060991304	SAINT PLACID'S HIGH SCHOOL	KOTWALI,CHITTAG NG-4000
021541060991307	KRISHNAKUMARI CITY CORPORATION. GIRL'S HIGH SCHOOL	KOTWALI,GPO
021541060991308	APARNA CHARAN CITY CORPORATION GIRL'S HIGH SCHOOL	KOTWALI,CHITTAGONG SADAR
021541060991309	GOVT. MUSLIM HIGH SCHOOL	KOTWALI,ANDAR KILLA

School ID	Schools Name	Thana
021541088951301	ST. SCHOLASTICA'S GIRL'S HIGH SCHOOL	KOTWALI,CHITTAGONG-4000
021541088951302	PATHERGHATA GIRL'S HIGH SCHOOL	KOTWALI,SADAR
021541088951303	BALUARDIGHI CITY CORP. GIRL'S HIGH SCHOOL	KOTWALI,CHITTAGONG-4000
021555211501302	HAJI ABDUL ALI CITY CORPORATION HIGH SCHOOL	PAHARTALI,FEROZSHAH
021555211501304	BANGLADEH RAILWAY GOVT. HIGH SCHOOL	PAHARTALI,PAHARTALI
021555217371303	SHAHID LT.G.M. MUSFIQUE BIRUTTAM HIGH SCHOOL	PAHARTALI,HOUSING ESTATE
021555223101302	FEROZ SHAH GIRL'S HIGH SCHOOL	PAHARTALI,FEROZ SHAH COLONY
021555229581203	MIRZA AHMED ISPAHANI SCHOOL	PAHARTALI,PAHARTALI
021555350361304	RABEYA BASRI GIRL'S HIGH SCHOOL	PAHARTALI,HALISHAHAR HOUSING
021557115631303	CHITTAGONG GOVT.GIRL'S HIGH SCHOOL	PANCHLAISH,KHULSI
021557115631305	AUNKUR SOCIETY GIRLS HIGH SCHOOL	PANCHLAISH,PANCHLAISH
021557115633005	ISPAHANI PUBLIC SCHOOL & COLLEGE	PANCHLAISH,G.P O
021557119951306	NASIRABAD GOVT. HIGH SCHOOL	PANCHLAISH,G.P.O CHITTAGONG
021557128211202	SHOLASHAHAR PUBLIC SCHOOL	PANCHLAISH,AMIN JUTE MILLS
021557143263003	CHITTAGONG PUBLIC HIGH SCHOOL & COLLEGE	BAYEJID BOSTAM,BAIZID BOSTAMI

Division: DHAKA
District: DHAKA

School ID	Schools Name	Thana
032604172001301	VIQARUNNISA NOON SCHOOL (BASUNDHARA BRANCH)	BADDA, BARIDHARA
032608010781308	BHASAN TEK HIGH SCHOOL	BHASHAN TEK , KAFRUL, DHAKA CANT.
032608012611303	DHAMALKOT ADARSHA ACADEMY	KAFRULDHAKA CANTONMENT
032608014951304	ADAMJEE CANTONMENT PUBLIC SCHOOL	CANTONMENT, DHAKA CANTONMENT
032608014953001	SHAHEED ANWAR GIRLS SCHOOLDHAKA	CANTONMENT, DHAKA CANTT.
032608754971303	KURMITOLA HIGH SCHOOL	BADDAK, HILKHET
032608759721305	UTTARA GIRL'S HIGH SCHOOL	UTTARA CANTONMENT, UTTARA
032608989981304	ADARSHA BIDYA NIKETAN	CANTONMENT, DHAKA CANTONMENT

School ID	Schools Name	Thana
032608989981310	SHAHID RAMIZ UDDIN CANTMENT SCHOOL	CANTONMENT
032608989983001	B A F SHAHEEN SCHOOL AND COLLEGE	DHAKA CANTT.
032608989983104	B N COLLEGE	DHAKA CANTONMENT
032612510531303	CITY CORPORATION ADARSHA HIGH SCHOOL	DEMRA,WARI
032612512421301	TOMIZUDDIN HIGH SCHOOL	DEMRA,WARI
032616174841301	RAYER BAZAR HIGH SCHOOL	DHANMONDI,RAYER BAZAR
032616182711301	KAKOLI HIGH SCHOOL	DHANMONDI,JIGATALA
032616182711304	KAMRUNNESSA GOVT GIRL'S HIGH SCHOOL	DHANMONDI ,ZIGATOLA
032616182711305	DHANMONDI GOVT GIRL'S HIGH SCHOOL	DHANMONDI,ZIGATOLA
032616182711306	DHANMONDI GOVT BOYS HIGH SCHOOL	DHANMONDI,MOHAMMADPUR
032616182711307	JUNIOR LABORATORY HIGH SCHOOL	DHANMONDI,ZIGATOLA
032616182711308	VIQARUNNISA NOON SCHOOL	DHANMONDI,DHANMONDI
032616182713110	Y.W.C.A HIGHER SECONDARY GIRLS SCHOOL	DHANMONDI,NEW MARKET
032616198601301	LAKE CIRCUS GIRLS HIGH SCHOOL	DHANMONDI,NEW MARKET
032616203251302	GOVERNMENT LABORATORY HIGH SCHOOL	DHANMONDI,DHANMONDI
032616206841301	BCSIR HIGH SCHOOL DHAMNONDI	DHANMONDI,NEW MARKE
032626589953102	SOUTH POINT SCHOOL & COLLEGE	GULSHAN ,GULSHAN
032626720551301	BANANI BIDYANIKETON	GULSHAN, BANANI
032626720551308	BANANI MODEL SCHOOL	GULSHAN, BANANI
032626722941306	GULSHAN MODEL HIGH SCHOOL	GULSHAN,GULSHAN
032626723651302	RAMPURA EKRAMUNNESSA HIGH SCHOOL	GULSHAN ,KHILGAON
032626727681309	MOHAKHALI MODEL HIGH SCHOOL	GULSHAN, GULSHAN
032626743151303	BADDA ALATUNNESA HIGH SCHOOL	BADDA, GULSHAN
032640315683001	AHMED BAWANY ACADEMY (SCHOOL & COLLEGE)	KOTWALI ,DHAKA SADAR
032640315683003	AHMED BAWANY ACADEMY	KOTWALI, DHAKA SADAR
032642215743002	BIR SHRESHTHA NOOR MOHAMMAD RIFLES PUBLIC SCHOOL & COLLEGE	LALBAGH,NEWMARKET
032642220561301	BIR SHRESHTHA MUNSHI ABDUR ROUF RIFLES COLLEGE	LALBAGH,NEWMARKET
032642240581301	AGRANI SCHOOL & COLLEGE	LALBAGH, NEWMARKET
032642240583001	AZIMPUR GIRLS SCHOOL & COLLEGE	LALBAGH ,NEWMARKET
032642252951303	ENGINEERING UNIVERSITY GIRLS' SCHOOL	LALBAGH,NEWMARKET
032648016333001	SOS HERMANN MEINER COLELGE	MIRPUR,MIRPUR-10
032648042621303	CANTONMENT BOARD HIGH SCHOOL	PALLABI,MIRPUR CANTONMENT
032648059251303	MONIPUR HIGH SCHOOL	PALLABI,MIRPUR
032650110461304	ARAB MISSION PUBLIC SCHOOL	MOHAMMADPUR

School ID	Schools Name	Thana
032650142103101	MOHAMMADPUR MODEL SCHOOL	MOHAMMADPUR
032650142351307	MOHAMMADPUR GOVT HIGH SCHOOL	MOHAMMADPUR
032650144393104	LALMATIA HOUSING SOCIETY HIGHER SECONDARY SCHOOL	MOHAMMADPUR,MOHAMMADPUR
032654525233002	MOTIJHEEL MODEL HIGH SCHOOL & COLLEGE	PALTAN, MOTIJHEEL
032654545421304	MOTIJHEEL GOVT BOYS HIGH SCHOOL	PALTAN, G. P. O
032654545421305	MOTIJHEEL GOVT GIRLS HIGH SCHOOL	PALTAN, SANTI NAGAR
032666639611303	SEGUN BAGICHA HIGH SCHOOL	RAMNA,G.P.O.
032666646101303	NILKHET HIGH SCHOOL	RAMNA ,G.P.O
032666646103001	UDAYAN HIGH SCHOOL	RAMNA,G.P.O.
032666646103002	ENGINEERING UNIVERSITY SCHOOL & COLLEGE	RAMNA, RAMNA
032666647983001	UNIVERSITY LABORATORY SCHOOL AND COLLEGE	RAMNA,RAMNA
032666647983004	UNIVERSITY LABORATORY HIGH SCHOOL	RAMNA,RAMNA
032666650101307	WILLES LITTLE FLOWER SCHOOLAND COLLEGE	RAMNA,G.P.O.
032666653543004	NATIONAL BANK PUBLIC HIGH SCHOOL	RAMNA, SANTINAGOR
032666673543001	NATIONAL BANK PUBLIC SCHOOL AND COLLEGE	RAMNA, SANTI NAGAR
032666673543002	ISPAHANI GIRLS SCHOOL AND COLLEGE	RAMNA, RAMNA
032668516471301	KHILGAON GOVT HIGH SCHOOL	SABUJBAGH, KHILGAON
032668572611304	SHANTIBAG HIGH SCHOOL	SABUJBAGH, SHANTINAGAR
032668572613001	KHILGON GIRLS SCHOOL AND COLLEGE	SABUJBAGH, KHILGAON
032668572613005	KHILGAON GIRLS HIGH SCHOOL AND COLLEGE	SABUJBAGH, KHILGAON
032690689041306	TEJGAON GOVT. GIRL'S HIGH SCHOOL	TEJGAON ,TEJGAON
032690689041307	TEJGAON GOVT. HIGH SCHOOL	TEJGAON, TEJGAON
032690689961303	BOTTOMLEY HOME GIRLS HIGH SCHOOL	TEJGAON ,TEJGOAN
032690689961305	GOVT. SCIENCE COLLEGE (SCHOOL SECTION)	TEJGAON ,TEJGAON
032690689961306	TEJGAON MODEL HIGH SCHOOL	TEJGAON, TEJKUNIPARA
032690695901302	RAJDHANI UCHCHA BIDDALAYA	TEJGAON, MOHAMMAD PUR
032690695901303	SHER-E-BANGLA NAGAR GOVT. GIRLS HIGH SCHOOL	TEJGAON,MOHAMMADPUR
032695384003005	RAJUK UTTARA MODEL SCHOOL & COLLEGE	UTTARA, UTTARA

Division: KHULNA

District: KHULNA

Institute Id	Institute Name	Thana, Post Office
044751014781301	HAZI ABDUL MALEK GIRLS HIGH SCHOOL	KHULNA SADAR,SHIPYARD
044751017741303	SHIPYARD SECONDARY SCHOOL	KHULNA SADAR,SHIPYARD
044751019941202	HAZI ABDUL MALAQUE JUNIOR HIGH SCHOOL	KHULNA SADAR,SHIPYARD
044751029181302	DHAKA MATCH INDUSTRIES SECONDARY SCHOOL	KHULNA SADAR,SHIPYARD
044751039083001	KHULNA COLLEGIATE GIRLS SCHOOL	KHULNA SADAR,KHULNA
044751039751201	WEST TOOT PARA JUNIOR SCHOOL	KHULNA SADAR,KHULNA
044751040281303	SULTANA HAMID ALI GIRLS HIGH SCHOOL	KHULNA SADAR,KHULNA
044751042961301	SABURAN NESSA GIRL'S SCHOOL	KHULNA SADAR,KHULNA
044751047071302	RUPSA MULTILATERAL HIGH SCHOOL	KHULNA SADAR,KHULNA
044751050281303	MODEL (MULTILATERAL) HIGH SCOOL	KHULNA SADAR,KHULNA
044751052961204	CORONATION SECONDARY BIDDYANICATON	KHULNA SADAR,KHULNA
044751052961302	FATIMA HIGH SCHOOL	KHULNA SADAR,KHULNA
044751056311301	PIONEER GIRLS HIGH SCHOOL	KHULNA SADAR,KHULNA
044751059851306	KHULNA ZILLA SCHOOL	KHULNA,SADAR KHULNA
044751060091201	KHULNA SHISHU HIGH SCHOOL	KHULNA SADAR,KHULNA
044751060091304	SAINT JOSEPH'S HIGH SCHOOL	KHULNA SADAR,KHULNA
044751060091305	GOVT. CORONATION GIRLS HIGH SCHOOL	KHULNA SADAR,KHULNA
044751061531302	A.P.C. GIRL'S HIGH SCHOOL	KHULNA SADAR,KHULNA
044751073631301	UDAYAN KHULNA ZILLA POLICE SCHOOL	KHULNA SADAR,KHULNA
044751077841302	KHANJAHAN ALI NIGHT HIGH SCHOOL	KHULNA SADAR,KHULNA
044751086601302	SHAHID SUHRAWARDY HIGH SCHOOL	KHULNA SADAR,KHULNA-9100
044751086601303	SHAHID SHAHRAWARDHY GIRLS HIGH SCOOL	KHULNA SADAR,KHULNA
044751086641301	HANNAY RAILWARY HIGH SCHOOL	KHULNA SADAR,KHULNA
044751086641305	RAILWAY GIRLS HIGH SCHOOL	KHULNA SADAR,KHULNA
044751090381302	SHAHID ZIA SECONDARY GIRLS HIGH SCHOOL	KHULNA SADAR,KHULNA
044751092291201	NIRALA ADARSHA HIGH SCHOOL	KHULNA SADAR,KHULNA CITY
044751103061303	SOS HERMANN GMEINER SCHOOL	KHULNA SADAR,KHULNA
044751103531301	SONAPOTA HIGH SCHOOL	KHULNA SADAR,KHULNA
044751103531302	IQBAL NAGAR GIRLS HIGH SCOOL	KHULNA SADAR,KHULNA

Division: RAJSHAHI

District: RAJSHAHI

Institute Id	Institute Name	Thana, Post Office
058122113751301	RAJSHAHI MUSLIM HIGH SCHOOL	BOALIA (SADAR),RAJSHAHI-6000
058122113751302	RAJSHAHI BAHUMUKHI GIRL'S HIGH SCHOOL	BOALIA (SADAR),RAJSHAHI
058122120621302	SHAHID NAZMUL HAQ GIRL'S HIGH SCHOOL	BOALIA (SADAR),RAJSHAHI
058122125811303	GOVT.P.N. GIRLS HIGH SCHOOL	BOALIA(SADAR),GHORAMARA
058122128263001	MASJID MISSION ACADEMY	BOALIA(SADAR),GHORAMARA
058122138631201	MOAZID MISSION ACADEMY	BOALIA(SADAR),BINODPUR BAZAR
058122147321301	RANI BAZAR GIRL'S HIGH SCHOOL	BOALIA(SADAR),GHORAMARA
058122174411301	ATKOSHI HIGH SCHOOL	BOALIA(SADAR),SUPURA
058122178951302	HAMIDPUR NAODA PARA HIGH SCHOOL	BOALIA(SADAR),SAPURA
058122178951303	RAJSHAHI SATELLITE TOWN HIGH SCHOOL	BOALIA(SADAR),SAPURA
058122178951304	NAODAPARA GIRL'S HIGH SCHOOL	BOALIA(SADAR),SAPURA
058122178951305	RAJSHAHI CANTONMENT BOARD HIGH SCHOOL	BOALIA(SADAR),RAJSHAHI CANTANMENT
058122178951306	SHAH MUKDUM HIGH SCHOOL	BOALIA(SADAR),SAPURA
058122178951307	HOUSIN ESTATE GIRLS HIGH SCHOOL	BOALIA(SADAR),RAJ.CANTON MENT
058122202541205	NIMNA MADHAMIC ADARSHAY GIRL'S SCHOOL	BOALIA(SADAR),SAPURA 11
058122202541302	SEROIL COLONY HIGH SCHOOL	BOALIA(SADAR),GHORAMARA
058122202541306	CHHOTO BONGRAM MADDYAMIC ADARSHA BALIKA BIDDYALOY	BOALIA(SADAR),SAPURA
058122209081303	SURJAKANA HIGH SCHOOL	BOALIA(SADAR),GHORAMARA
058122209081304	SEROIL GOVT. HIGH SCHOOL	BOALIA(SADAR),GHORAMARA
05812220995130	BALIA PUKUR VIDDYA NIKETON	BOALIA(SADAR),GHORAMARA
05812221168130	RAJSHAHI NIGHT HIGH SCHOOL	BOALIA(SADAR),RAJSHAHI
058122211681302	RAJSHAHI LOKNATH HIGH SCHOOL	BOALIA(SADAR),RAJSHAHI
058122211681303	RAJSHAHI COLLEGIATE SCHOOL	BOALIA (SADAR),BOALIA
058122246751301	RAJSHAHI ADARSHA UCHCHA BIDAYLAY	BOALIA(SADAR),KAZLA
058122246751304	SABITRI GIRL'S HIGH SCHOOL	BOALIA(SADAR),GHORAMARA
058122246753002	KHADEMUL ISLAM GIRL'S SCHOOL & COLLEGE	BOALIA(SADAR),GHORAMARA
058122267501301	MEHER CHONDI HIGH SCHOOL	BOALIA(SADAR),PADMA RESIDENSIAL
058122271613001	AGRANI SCHOOL & COLLEGE	BOALIA(SADAR),KAJLA
058122282111201	SAYRA KHATUN JUNIOR GIRL'S SCHOOL	BOALIA(SADAR),BENODPUR BAZAR
058122284111201	SUMMIT SCHOOL	BOALIA(SADAR),KAJLA
058122289513001	RAJSHAHI UNIVERSITY SCHOOL	MATIHAR,RAJ.UNIVERSITY

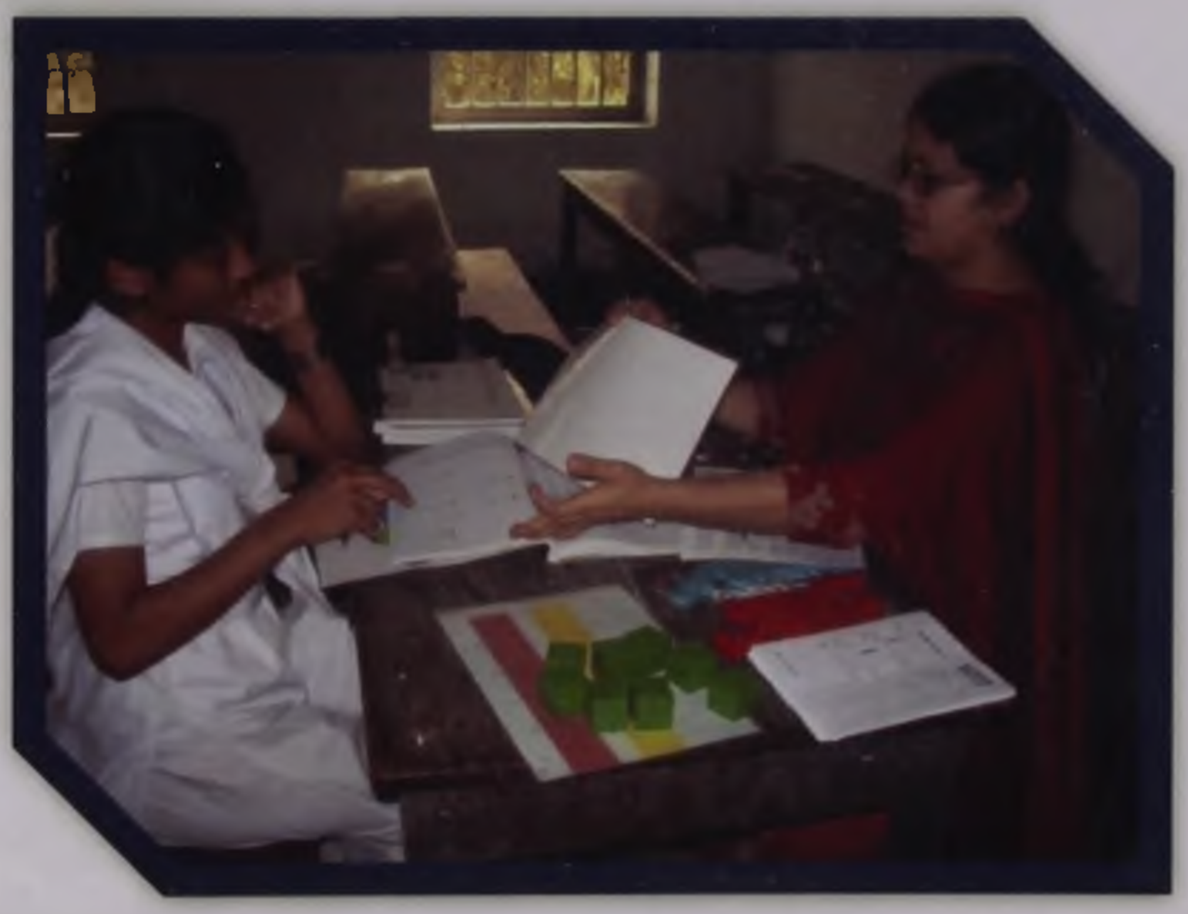
Institute Id	Institute Name	Thana, Post Office
058122298811302	DANSHMARI HIGH SCHOOL	BOALIA(SADAR),BINODPUR BAZAR
058122305111302	SHAHID ZIAUR RAHMAN HIGH SCHOOL	MATIHAR,KHORKHORI
058122305611303	MIRJAPUR HIGH SCHOOL	BOALIA(SADAR),BENODPUR BAZAR
058122305611304	B.C.S.I.R. LABORATORY HIGH SCHOOL	BOALIA(SADAR),BINODPUR BAZAR

Division: SYLHET
District: SYLHET

Institute Id	Institute Name	Thana, Post Office
069162012161301	BLUE BIRD HIGH SCHOOL	KOTWALI,SYLHET
069162012201302	RASAMOY MEMORIAL HIGH SCHOOL	KOTWALI,SYLHET
069162017581303	POLICE LINE HIGH SCHOOL	KOTWALI,SYLHET
069162028761301	MOINUN-NESSA GIRLS HIGH SCHOOL	KOTWALI,SYLHET
069162034991303	GOVT. AGRAGAMI GIRLS HIGH SCHOOL	KOTWALI,SYLHET
069162034991304	SYLHET GOVT. PILOT HIGH SCHOOL	KOTWALI,SAYLHET
069162036301201	MIRZA JANGAL JR GIRLS HIGH SCHOOL	KOTWALI,SYLHET
069162036891302	RAJA G.C. HIGH SCHOOL	KOTWALI,SYLHET-3100
069162040281301	AMBER KHANA GIRL'S SCHOOL	KOTWALI,SYEHET
069162040283002	AMBORKHANA GIRLS SCHOOL AND COLLEGE	KOTWALI,SYLHET-3100
069162044921302	MODEL HIGH SCHOOL	KOTWALI,SYLHET
069162044921303	KAZI JALAL UDDIN GIRLS ML HIGH SCHOOL	KOTWALI,SYLHET
069162047191305	KISHORI MOHAN GIRLS HIGH SCHOOL	KOTWALI,SYLHET
069162049261304	THE AIDED HIGH SCHOOL	KOTWALI,SYLHET
069162051541301	RAMKRISHNA GIRLS HIGH SCHOOL	KOTWALI,SYLHET-3100
069162053403001	SHAHJALAL JAMIA ISLAMIA SCHOOL & COLLEGE	KOTWALI,SYLHET
069162059981303	SHAHJALAL UPASHAHAR HIGH SCHOOL	KOTWALI,UPO SHAHAR
069162651133003	MOHAMMAD MOKON HIGH SCHOOL	KOTWALI,SYLHET
069162653211301	JALALABAD BL.HIGH SCHOOL	KOTWALI,SYLHET
069162655641302	HAZI RASHID ALI HIGH SCHOOL	KOTWALI,KAMAL BAZAR
069162900981201	BIRESH CH0NDRA JUNIOR HIGH SCHOOL	KOTWALI,AKHALIA
069162904451303	RIFLES PUBLIC SCHOOL	KOTWALI,AKHALIA
069162905441302	HAJI ABDUS SATTAR HIGH SCHOOL	KOTWALI,TUKERBAZAR
069162905541301	SHAHAJALAL UNIVERSITY HIGH SCHOOL	KOTWALI,SABIPRABI
069162905941201	MOERCHAR JUNIOR HIGH SCHOOL	KOTWAL,SONATAL
069162950681302	PATHANTULA BILATERAL HIGH SCHOOL	KOTWALI,SYLHET
069162950681303	PDB HIGH SCHOOL	KOTWALI,SYLHET

Institute Id	Institute Name	Thana, Post Office
06916295821130l	SYED HATIM ALI HIGH SCHOOL	KOTWALI,SYLHET
06916295985130‹	HAZI MOHAMMAD SHAFIQUE HIGH SCHOOL.	KOTWALI,SYLHET-3100

Appendix 8: Pictures of Test Administration



Appendix 9: Item wise Difficulty and Discrimination Index on All Age Group

Domain	Items	U_p	U_l	U	L	P	D	Answered
Nonverbal	KN_2_1	11	4	11	4	1.00	0.64	15
Nonverbal	KN_2_2	11	4	11	4	1.00	0.64	15
Nonverbal	KN_2_3	11	4	11	4	1.00	0.64	15
Nonverbal	KN_2_4	11	4	11	4	1.00	0.64	15
Nonverbal	KN_2_5	11	3	11	4	0.93	0.73	15
Nonverbal	KN_2_6	11	4	11	4	1.00	0.64	15
Nonverbal	KN_3_1	33	57	40	71	0.81	-0.60	111
Nonverbal	KN_3_2	24	28	40	71	0.47	-0.10	111
Nonverbal	KN_3_3	19	35	40	71	0.49	-0.40	111
Nonverbal	KN_3_4	37	58	40	71	0.86	-0.53	111
Nonverbal	KN_3_5	34	65	40	71	0.89	-0.78	111
Nonverbal	KN_3_6	29	43	40	71	0.65	-0.35	111
Nonverbal	KN_4_1	80	51	81	62	0.92	0.36	143
Nonverbal	KN_4_2	44	9	81	61	0.37	0.43	142
Nonverbal	KN_4_3	67	28	81	61	0.67	0.48	142
Nonverbal	KN_4_4	22	0	81	61	0.15	0.27	142
Nonverbal	KN_4_5	43	8	81	61	0.36	0.43	142
Nonverbal	KN_4_6	40	0	81	61	0.28	0.49	142
Nonverbal	KN_5_1	47	3	58	3	0.82	0.76	61
Nonverbal	KN_5_2	13	0	58	3	0.21	0.22	61
Nonverbal	KN_5_3	38	0	58	3	0.62	0.66	61
Nonverbal	KN_5_4	5	0	58	3	0.08	0.09	61
Nonverbal	KN_5_5	19	0	58	3	0.31	0.33	61
Nonverbal	KN_5_6	7	0	58	3	0.11	0.12	61
Nonverbal	KN_6_1	15	0	18	0	0.83	0.83	18
Nonverbal	KN_6_2	12	0	18	0	0.67	0.67	18
Nonverbal	KN_6_3	5	0	18	0	0.28	0.28	18
Nonverbal	KN_6_4	5	0	18	0	0.28	0.28	18
Nonverbal	KN_6_5	1	0	18	0	0.06	0.06	18
Nonverbal	KN_6_6	1	0	18	0	0.06	0.06	18
Nonverbal	NV_FR_1	0	0	0	0	0	0	0
Nonverbal	NV_FR_10	13	1	14	1	0.93	0.86	15
Nonverbal	NV_FR_11	13	1	14	1	0.93	0.86	15
Nonverbal	NV_FR_12	11	1	14	1	0.80	0.71	15
Nonverbal	NV_FR_13	39	45	50	55	0.80	-0.12	105
Nonverbal	NV_FR_14	35	49	39	58	0.87	-0.36	97
Nonverbal	NV_FR_15	29	49	39	58	0.80	-0.51	97
Nonverbal	NV_FR_16	28	44	39	58	0.74	-0.41	97
Nonverbal	NV_FR_17	49	63	64	76	0.80	-0.22	140
Nonverbal	NV_FR_18	30	20	30	21	0.98	0.33	51
Nonverbal	NV_FR_19	29	20	30	21	0.96	0.30	51
Nonverbal	NV_FR_2	0	0	0	0	0.00	0.00	0

Domain	Items	U_p	U_t	U	L	P	D	Answered
Nonverbal	NV_FR_20	29	20	30	20	0.98	0.30	50
Nonverbal	NV_FR_21	25	17	29	20	0.86	0.28	50
Nonverbal	NV_FR_22	25	17	30	20	0.84	0.27	50
Nonverbal	NV_FR_23	20	14	30	18	0.71	0.20	48
Nonverbal	NV_FR_24	16	14	26	18	0.68	0.08	44
Nonverbal	NV_FR_25	14	12	25	16	0.63	0.08	41
Nonverbal	NV_FR_26	14	11	23	15	0.66	0.13	38
Nonverbal	NV_FR_27	9	9	18	15	0.55	0.00	33
Nonverbal	NV_FR_28	5	8	17	14	0.42	-0.18	31
Nonverbal	NV_FR_29	4	3	13	12	0.28	0.08	25
Nonverbal	NV_FR_3	0	0	0	0	0.00	0.00	0
Nonverbal	NV_FR_30	4	2	11	11	0.27	0.18	22
Nonverbal	NV_FR_31	4	3	9	10	0.37	0.11	19
Nonverbal	NV_FR_32	1	1	7	5	0.17	0.00	12
Nonverbal	NV_FR_33	1	0	6	3	0.11	0.17	9
Nonverbal	NV_FR_34	0	2	5	2	0.29	-0.40	7
Nonverbal	NV_FR_35	0	0	1	2	0.00	0.00	3
Nonverbal	NV_FR_36	0	1	0	2	0.50	0.00	2
Nonverbal	NV_FR_4	13	1	13	1	1.00	0.92	14
Nonverbal	NV_FR_5	12	1	14	1	0.87	0.79	15
Nonverbal	NV_FR_6	12	1	14	1	0.87	0.79	15
Nonverbal	NV_FR_7	11	1	14	1	0.80	0.71	15
Nonverbal	NV_FR_8	12	1	14	1	0.87	0.79	15
Nonverbal	NV_FR_9	13	1	14	1	0.93	0.86	15
Nonverbal	QR_2_1	11	4	11	4	1.00	0.64	15
Nonverbal	QR_2_2	11	4	11	4	1.00	0.64	15
Nonverbal	QR_2_3	11	4	11	4	1.00	0.64	15
Nonverbal	QR_2_4	11	4	11	4	1.00	0.64	15
Nonverbal	QR_2_5	11	4	11	4	1.00	0.64	15
Nonverbal	QR_2_6	11	4	11	4	1.00	0.64	15
Nonverbal	QR_3_1	40	79	40	79	1.00	-0.98	119
Nonverbal	QR_3_2	40	79	40	79	1.00	-0.98	119
Nonverbal	QR_3_3	40	79	40	79	1.00	-0.98	119
Nonverbal	QR_3_4		79	40	79	1.00	-0.98	119
Nonverbal	QR_3_5	40	79	40	79	1.00	-0.98	119
Nonverbal	QR_3_6	40	79	40	79	1.00	-0.98	119
Nonverbal	QR_4_1	77	77	81	81	0.95	0.00	162
Nonverbal	QR_4_2	47	22	81	81	0.43	0.31	162
Nonverbal	QR_4_3	81	76	81	81	0.97	0.06	162
Nonverbal	QR_4_4	55	22	81	81	0.48	0.41	162
Nonverbal	QR_4_5	66	13	81	81	0.49	0.65	162
Nonverbal	QR_4_6	55	6	81	81	0.38	0.60	162
Nonverbal	QR_5_1	77	21	78	22	0.98	0.72	100
Nonverbal	QR_5_2	63	8	78	22	0.71	0.71	100
Nonverbal	QR_5_3	28	2	78	22	0.30	0.33	100

Domain	Items	U_p	U_l	U	L	P	D	Answered
Nonverbal	QR_5_4	24	0	78	22	0.24	0.31	100
Nonverbal	QR_5_5	15	0	78	22	0.15	0.19	100
Nonverbal	QR_5_6	8	0	78	22	0.08	0.10	100
Nonverbal	QR_6_1	35	0	37	0	0.95	0.95	37
Nonverbal	QR_6_2	25	0	37	0	0.68	0.68	37
Nonverbal	QR_6_3	8	0	37	0	0.22	0.22	37
Nonverbal	QR_6_4	8	0	37	0	0.22	0.22	37
Nonverbal	QR_6_5	5	0	37	0	0.14	0.14	37
Nonverbal	QR_6_6	2	0	36	0	0.06	0.06	36
Nonverbal	VS_2_1	11	4	11	4	1.00	0.64	15
Nonverbal	VS_2_2	10	3	11	4	0.87	0.64	15
Nonverbal	VS_2_3	9	0	11	4	0.60	0.82	15
Nonverbal	VS_2_4	8	0	11	4	0.53	0.73	15
Nonverbal	VS_2_5	4	0	11	4	0.27	0.36	15
Nonverbal	VS_2_6	0	0	11	4	0.00	0.00	15
Nonverbal	VS_3_1	37	52	40	77	0.76	-0.38	117
Nonverbal	VS_3_2	38	57	40	77	0.81	-0.48	117
Nonverbal	VS_3_3	35	32	40	77	0.57	0.08	117
Nonverbal	VS_4_1	74	42	81	43	0.94	0.40	155
Nonverbal	VS_4_2	50	33	81	69	0.55	0.21	155
Nonverbal	VS_4_3	16	3	81	74	0.12	0.16	155
Nonverbal	VS_5_1	24	0	24	0	1.00	1.00	38
Nonverbal	VS_5_2	23	0	33	0	0.70	0.70	38
Nonverbal	VS_5_3	6	0	38	0	0.16	0.16	38
Nonverbal	VS_6_1	13	0	13	0	1.00	1.00	17
Nonverbal	VS_6_2	11	0	16	0	0.69	0.69	17
Nonverbal	VS_6_3	2	0	17	0	0.12	0.12	17
Nonverbal	WM_2_1	11	4	11	4	1.00	0.64	15
Nonverbal	WM_2_2	11	4	11	4	1.00	0.64	15
Nonverbal	WM_2_3	11	4	11	4	1.00	0.64	15
Nonverbal	WM_2_4	10	4	11	4	0.93	0.55	15
Nonverbal	WM_2_5	11	1	11	4	0.80	0.91	15
Nonverbal	WM_2_6	10	3	11	4	0.87	0.64	15
Nonverbal	WM_3_1	39	71	40	79	0.92	-0.80	119
Nonverbal	WM_3_2	40	78	40	79	0.99	-0.95	119
Nonverbal	WM_3_3	40	75	40	79	0.97	-0.88	119
Nonverbal	WM_3_4	38	59	40	79	0.82	-0.53	119
Nonverbal	WM_3_5	39	74	40	79	0.95	-0.88	119
Nonverbal	WM_3_6	32	67	40	79	0.83	-0.88	119
Nonverbal	WM_4_1	75	67	80	75	0.92	0.10	155
Nonverbal	WM_4_2	73	51	80	75	0.80	0.28	155
Nonverbal	WM_4_3	59	44	80	75	0.66	0.19	155
Nonverbal	WM_4_4	59	23	80	75	0.53	0.45	155
Nonverbal	WM_4_5	56	7	80	75	0.41	0.61	155
Nonverbal	WM_4_6	43	3	80	75	0.30	0.50	155

Domain	Items	U_p	U_l	U	L	P	D	Answered
Nonverbal	WM_5_1	71	31	72	33	0.97	0.56	105
Nonverbal	WM_5_2	55	15	72	33	0.67	0.56	105
Nonverbal	WM_5_3	48	4	72	33	0.50	0.61	105
Nonverbal	WM_5_4	39	1	72	33	0.38	0.53	105
Nonverbal	WM_5_5	17	1	72	33	0.17	0.22	105
Nonverbal	WM_5_6	14	0	72	33	0.13	0.19	105
Nonverbal	WM_6_1	54	2	56	2	0.97	0.93	58
Nonverbal	WM_6_2	45	0	56	2	0.78	0.80	58
Nonverbal	WM_6_3	28	1	56	2	0.50	0.48	58
Nonverbal	WM_6_4	11	0	55	2	0.19	0.20	57
Nonverbal	WM_6_5	9	0	55	2	0.16	0.16	57
Nonverbal	WM_6_6	3	0	55	2	0.05	0.05	57
Verbal	FR_2_1	21	16	22	32	0.69	0.23	54
Verbal	FR_2_2	12	25	22	32	0.69	-0.59	54
Verbal	FR_2_3	11	25	22	32	0.67	-0.64	54
Verbal	FR_3_1	80	79	81	83	0.97	0.01	164
Verbal	FR_3_2	71	47	81	83	0.72	0.30	164
Verbal	FR_3_3	75	72	81	83	0.90	0.04	164
Verbal	FR_4_1	20	50	20	51	0.99	-1.50	143
Verbal	FR_4_2	44	33	46	57	0.75	0.24	143
Verbal	FR_4_3	42	13	70	62	0.42	0.41	143
Verbal	FR_5_1	37	1	38	1	0.97	0.95	43
Verbal	FR_5_2	23	0	41	1	0.55	0.56	43
Verbal	FR_5_3	8	0	42	1	0.19	0.19	43
Verbal	FR_6_1	5	0	5	0	1.00	1.00	6
Verbal	FR_6_2	4	0	6	0	0.67	0.67	6
Verbal	FR_6_3	2	0	6	0	0.33	0.33	6
Verbal	QR_2_10	19	28	19	28	1.00	-0.47	47
Verbal	QR_2_20	19	28	19	28	1.00	-0.47	47
Verbal	QR_2_30	19	28	19	28	1.00	-0.47	47
Verbal	QR_2_40	19	28	19	28	1.00	-0.47	47
Verbal	QR_2_50	19	28	19	28	1.00	-0.47	47
Verbal	QR_2_60	19	28	19	28	1.00	-0.47	47
Verbal	QR_3_10	81	82	81	83	0.99	-0.01	164
Verbal	QR_3_20	81	83	81	83	1.00	-0.02	164
Verbal	QR_3_30	81	82	81	83	0.99	-0.01	164
Verbal	QR_3_40	81	83	81	83	1.00	-0.02	164
Verbal	QR_3_50	81	80	81	83	0.98	0.01	164
Verbal	QR_3_60	81	75	81	83	0.95	0.07	164
Verbal	QR_4_10	35	13	67	15	0.59	0.33	82
Verbal	QR_4_20	30	7	67	15	0.45	0.34	82
Verbal	QR_4_30	48	9	67	15	0.70	0.58	82
Verbal	QR_4_40	50	2	67	15	0.63	0.72	82
Verbal	QR_4_50	39	1	67	15	0.49	0.57	82
Verbal	QR_4_60	24	0	67	15	0.29	0.36	82
Verbal	QR_5_10	26	0	50	0	0.52	0.52	50
Verbal	QR_5_20	31	0	50	0	0.62	0.62	50

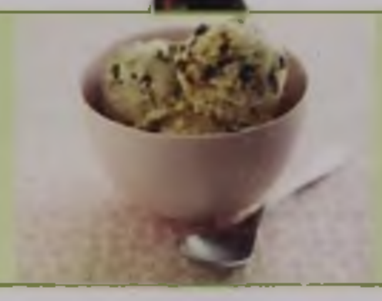
Domain	Items	U_p	U_l	U	L	P	D	Answered
Verbal	QR_5_30	33	0	50	0	0.66	0.66	50
Verbal	QR_5_40	16	0	50	0	0.32	0.32	50
Verbal	QR_5_50	15	0	50	0	0.30	0.30	50
Verbal	QR_5_60	6	0	50	0	0.12	0.12	50
Verbal	QR_6_10	24	0	25	0	0.96	0.96	25
Verbal	QR_6_20	14	0	25	0	0.56	0.56	25
Verbal	QR_6_30	9	0	25	0	0.36	0.36	25
Verbal	QR_6_40	12	0	25	0	0.48	0.48	25
Verbal	QR_6_50	0	0	25	0	0.00	0.00	25
Verbal	QR_6_60	0	0	25	0	0.00	0.00	25
Verbal	KN_1	0	0	0	0	0.00	0.00	0
Verbal	KN_10	35	29	35	29	1.00	0.17	64
Verbal	KN_11	35	29	35	29	1.00	0.17	64
Verbal	KN_12	35	29	35	29	1.00	0.17	64
Verbal	KN_13	31	26	35	29	0.89	0.14	64
Verbal	KN_14	79	78	81	78	0.99	0.01	159
Verbal	KN_15	22	15	23	14	1.00	0.30	95
Verbal	KN_16	29	35	29	36	0.98	-0.21	95
Verbal	KN_17	29	38	31	38	0.97	-0.29	95
Verbal	KN_18	17	17	19	19	0.89	0.00	95
Verbal	KN_19	23	30	30	32	0.85	-0.23	95
Verbal	KN_2	0	0	0	0	0.00	0.00	0
Verbal	KN_20	27	18	27	21	0.94	0.33	95
Verbal	KN_21	0	0	0	0	0	0	0
Verbal	KN_22	0	0	0	0	0	0	0
Verbal	KN_23	0	0	0	0	0	0	0
Verbal	KN_24	0	0	0	0	0	0	0
Verbal	KN_25	0	0	0	0	0	0	0
Verbal	KN_26	0	0	0	0	0	0	0
Verbal	KN_27	0	0	0	0	0	0	0
Verbal	KN_28	0	0	0	0	0	0	0
Verbal	KN_29	0	0	0	0	0	0	0
Verbal	KN_3	0	0	0	0	0	0	0
Verbal	KN_30	0	0	0	1	0	0	1
Verbal	KN_31	0	1	0	1	1	0	1
Verbal	KN_32	0	0	0	1	0	0	1
Verbal	KN_33	0	0	0	1	0	0	1
Verbal	KN_34	0	1	0	1	1	0	1
Verbal	KN_35	0	0	0	1	0	0	1
Verbal	KN_36	0	0	0	1	0	0	1
Verbal	KN_37	0	0	0	0	0	0	0
Verbal	KN_38	0	0	0	0	0	0	0
Verbal	KN_39	0	0	0	0	0	0	0
Verbal	KN_4	0	0	0	0	0	0	0
Verbal	KN_40	0	0	0	0	0	0	0
Verbal	KN_41	0	0	0	0	0	0	0
Verbal	KN_42	0	0	0	0	0	0	0
Verbal	KN_43	0	0	0	0	0	0	0

Domain	Items	U_p	U_l	U	L	P	D	Answered
Verbal	KN_44	0	0	0	0	0	0	0
Verbal	KN_5	0	0	0	0	0	0	0
Verbal	KN_6	0	0	0	0	0	0	0
Verbal	KN_7	0	0	0	0	0	0	0
Verbal	KN_8	0	1	0	1	1	0	1
Verbal	KN_9	33	28	33	29	0.98	0.15	62
Verbal	VS_2_10	19	28	19	28	1.00	-0.47	47
Verbal	VS_2_20	18	28	19	28	0.98	-0.53	47
Verbal	VS_2_30	17	26	19	28	0.91	-0.47	47
Verbal	VS_2_40	18	25	19	28	0.91	-0.37	47
Verbal	VS_2_50	19	28	19	28	1.00	-0.47	47
Verbal	VS_2_60	19	28	19	28	1.00	-0.47	47
Verbal	VS_3_10	63	45	81	83	0.66	0.22	164
Verbal	VS_3_20	49	40	81	83	0.54	0.11	164
Verbal	VS_3_30	80	80	81	83	0.98	0.00	164
Verbal	VS_3_4	81	81	81	83	0.99	0.00	164
Verbal	VS_3_5	78	64	81	83	0.87	0.17	164
Verbal	VS_3_6	77	66	81	83	0.87	0.14	164
Verbal	VS_4_10	82	73	82	76	0.98	0.11	158
Verbal	VS_4_20	81	69	82	76	0.95	0.15	158
Verbal	VS_4_30	78	70	82	76	0.94	0.10	158
Verbal	VS_4_4	71	27	82	76	0.62	0.54	158
Verbal	VS_4_5	33	4	82	76	0.23	0.35	158
Verbal	VS_4_6	17	0	82	76	0.11	0.21	158
Verbal	VS_5_10	38	25	65	29	0.67	0.20	94
Verbal	VS_5_20	49	6	65	29	0.59	0.66	94
Verbal	VS_5_30	32	2	65	29	0.36	0.46	94
Verbal	VS_5_4	29	0	65	28	0.31	0.45	93
Verbal	VS_5_5	15	0	65	29	0.16	0.23	94
Verbal	VS_5_6	5	0	65	29	0.05	0.08	94
Verbal	VS_6_10	29	0	32	0	0.91	0.91	32
Verbal	VS_6_20	3	0	32	0	0.09	0.09	32
Verbal	VS_6_30	12	0	32	0	0.38	0.38	32
Verbal	VS_6_4	17	0	32	0	0.53	0.53	32
Verbal	VS_6_5	2	0	32	0	0.06	0.06	32
Verbal	VS_6_6	0	0	32	0	0.00	0.00	32
Verbal	WM_2_10	19	27	19	28	0.98	-0.42	47
Verbal	WM_2_20	18	24	19	28	0.89	-0.32	47
Verbal	WM_2_30	16	19	3	28	1.13	-1.00	47
Verbal	WM_3_10	79	60	80	83	0.85	0.24	164
Verbal	WM_3_20	60	36	80	83	0.59	0.30	164
Verbal	WM_3_30	36	17	80	83	0.33	0.24	164
Verbal	WM_4_10	69	16	81	51	0.64	0.65	132
Verbal	WM_4_20	54	9	81	51	0.48	0.56	132
Verbal	WM_4_30	42	8	81	51	0.38	0.42	132
Verbal	WM_5_10	35	11	36	13	0.94	0.67	92
Verbal	WM_5_20	48	9	51	15	0.86	0.76	92
Verbal	WM_5_30	54	5	67	17	0.70	0.73	92

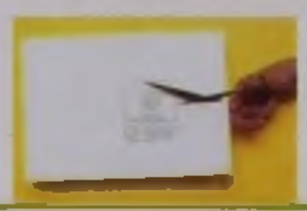
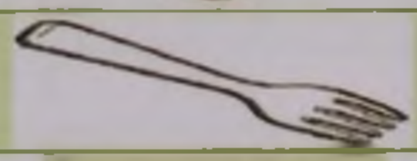
Domain	Items	U_p	U_l	U	L	P	D	Answered
Verbal	WM_6_10	43	2	43	2	1.00	0.95	58
Verbal	WM_6_20	50	2	51	2	0.98	0.94	58
Verbal	WM_6_30	45	1	53	2	0.84	0.83	58

Appendix 10: Pictorial Adapted Items (Non verbal and Verbal)

Pictorial Non- Verbal Adapted Items

Item No.	Original Item	Replaced Item	Reason/note	Adapted items
KN2_1	Feeds child	Theme unchanged	Replaced with local picture	
KN3_2	Drinks with straw from glass	Bottle with straw	Culturally appropriate picture introduced	
KN3_3	Eat with spoon from big bowl	Eat with spoon from a small bowl	Culturally appropriate picture introduced	
KN3_6	Sweeps with mop	Sweeps with Local broom	Culturally appropriate picture introduced	
KN4_2	Stamp in wrong place	Theme unchanged	Culturally appropriate picture introduced	
KN4_3	Balanced scales	Theme unchanged	Culturally appropriate picture introduced	
KN4_5	Rooster on nest	Theme unchanged	Culturally more clear picture introduced	
QR2_4	5 birds	Theme unchanged	Replaced with local birds	

Pictorial Verbal Adapted Items

Item No.	Original Item	Replaced Item	Reason/Note	Adapted items
Kn_10	Child Drinking	Theme unchanged	Culturally appropriate picture introduced	
Kn_11	Cutting a piece of paper	Theme unchanged	Culturally appropriate picture introduced	
Kn_12	Running	Boy running	Culturally appropriate picture introduced	
Kn_13	Tying shoe	Theme unchanged	Culturally appropriate picture introduced	
Kn_14	Writing on a paper	Theme unchanged	Culturally appropriate picture introduced	
FR2_1	Cat and ball playing	Theme unchanged	Culturally appropriate picture introduced	
FR2_2	Laundry	Theme unchanged	Culturally appropriate picture introduced	
FR2_3	Puzzle	Theme unchanged	Culturally appropriate picture introduced	
QR2_4	2 dogs	Theme unchanged	Culturally appropriate picture introduced	
FR3_3	Rugby ball	Cricket ball	Culturally appropriate picture introduced	
FR3_4	Chop-stick	Fork	Culturally appropriate picture introduced	
VS3_1	In front of girl	Rearranged as VS3_3	Difficult where VS3_3 was easier P=(0.98)	

Appendix 11: Tables for Norm

Table. *Mean and SD of Raw Scores- Male Student*

		Non Verbal					Verbal				
Age		RS_NV_FR	RS_NV_KN	RS_NV_QR	RS_NV_VS	RS_NV_WM	RS_V_FR	RS_V_KN	RS_V_QR	RS_V_VS	RS_V_WM
6	M	11.3	10.8	14.5	14.5	18.5	10.4	28.0	11.7	12.9	10.9
	SD	1.4	1.7	1.8	2.8	4.0	2.4	5.4	1.3	2.0	4.0
	N	150	150	150	150	150	150	150	150	150	150
7	M	15.3	12.1	14.3	16.8	18.7	10.7	29.1	11.7	13.1	10.6
	SD	1.7	1.8	1.5	2.2	2.4	2.7	5.9	1.5	2.3	3.8
	N	150	150	150	150	150	150	150	150	150	150
8	M	16.0	12.2	14.9	17.5	20.0	11.5	31.3	11.8	13.6	11.8
	SD	1.0	1.7	1.4	1.8	2.6	1.9	5.1	0.9	2.2	4.0
	N	150	150	150	150	150	150	150	150	150	150
9	M	15.9	13.1	15.1	17.1	19.6	11.6	31.4	11.8	13.6	11.8
	SD	1.0	2.0	1.4	2.7	2.7	2.3	5.9	1.5	2.1	4.1
	N	150	150	150	150	150	150	150	150	150	150
10	M	16.1	12.9	15.1	17.9	20.1	13.4	40.7	12.2	14.5	12.6
	SD	1.0	1.9	1.3	1.4	2.6	1.4	7.0	1.2	1.9	4.5
	N	150	150	150	150	150	150	150	150	150	150
11	M	16.4	13.4	15.6	18.2	20.9	13.9	42.6	12.3	15.0	13.7
	SD	1.4	2.4	1.4	1.3	2.8	1.8	6.7	0.9	1.7	4.7
	N	150	150	150	150	150	150	150	150	150	150
12	M	16.4	13.5	15.8	18.5	21.3	14.0	43.7	12.5	15.0	12.9
	SD	1.0	2.3	1.7	1.4	3.0	1.2	7.8	1.6	1.7	3.7
	N	150	150	150	150	150	150	150	150	150	150
13	M	23.4	15.3	17.2	19.3	22.9	14.4	47.5	12.6	15.6	14.2
	SD	2.8	3.0	2.3	1.8	3.0	1.5	4.8	1.5	2.1	4.0
	N	150	150	150	150	150	150	150	150	150	150
14	M	24.3	16.3	18.1	19.9	24.0	14.7	48.0	13.6	16.5	15.6
	SD	2.9	3.5	2.4	2.2	3.1	1.6	5.4	2.4	2.7	5.3
	N	150	150	150	150	150	150	150	150	150	150
15	M	25.2	17.2	19.2	20.9	25.4	15.6	50.2	14.6	17.5	17.4
	SD	2.7	4.1	3.4	2.7	3.5	1.9	7.1	3.4	2.9	4.5
	N	150	150	150	150	150	150	150	150	150	150
16	M	25.6	18.3	20.5	21.4	26.8	16.0	51.6	15.8	18.2	19.6
	SD	2.7	4.0	3.5	2.7	3.6	2.8	8.7	4.3	3.3	4.7
	N	150	150	150	150	150	150	150	150	150	150

Table

Mean and SD of Raw Scores- Female Student

		Non Verbal					Verbal				
Age		RS_NV_FR	RS_NV_KN	RS_NV_QR	RS_NV_VS	RS_NV_WM	RS_V_FR	RS_V_KN	RS_V_QR	RS_V_VS	RS_V_WM
6	M	11.6	10.5	14.2	13.8	18.4	10.2	26.7	11.3	12.6	10.2
	SD	1.3	1.6	1.8	3.1	3.4	2.2	5.6	1.7	2.3	3.5
	N	150	150	150	150	150	150	150	150	150	150
7	M	15.7	11.9	14.5	16.8	18.8	10.4	28.6	11.5	12.9	10.5
	SD	1.3	1.6	1.4	2.0	2.9	2.7	6.5	1.4	2.3	4.2
	N	150	150	150	150	150	150	150	150	150	150
8	M	15.9	12.3	14.6	17.2	18.8	10.8	29.6	11.8	13.3	11.2
	SD	1.1	1.9	1.8	2.2	2.6	2.6	5.8	1.3	2.6	4.0
	N	150	150	150	150	150	150	150	150	150	150
9	M	15.9	12.6	15.1	17.3	19.5	11.4	30.1	11.9	13.7	11.3
	SD	1.0	1.7	1.7	2.5	2.4	2.3	6.5	1.1	2.2	3.9
	N	150	150	150	150	150	150	150	150	150	150
10	M	16.2	12.8	15.2	17.9	20.0	13.3	41.0	12.1	14.4	12.0
	SD	0.9	1.5	1.3	1.3	2.6	1.1	6.5	0.7	1.7	3.7
	N	150	150	150	150	150	150	150	150	150	150
11	M	16.2	13.3	15.1	18.0	20.7	13.9	42.4	12.5	14.8	13.1
	SD	0.9	2.5	1.6	1.8	2.8	1.6	7.0	1.4	2.1	4.7
	N	150	150	150	150	150	150	150	150	150	150
12	M	16.5	13.2	15.5	18.2	20.6	13.9	43.9	12.4	15.2	13.3
	SD	1.5	2.3	1.3	1.8	2.5	1.1	6.1	0.9	1.7	4.1
	N	150	150	150	150	150	150	150	150	150	150
13	M	23.4	14.6	16.8	18.8	23.0	14.2	46.3	12.9	15.7	14.2
	SD	2.6	2.9	2.2	2.0	2.8	1.3	5.3	1.8	2.3	4.2
	N	150	150	150	150	150	150	150	150	150	150
14	M	24.2	15.7	18.1	19.6	24.5	14.8	47.8	13.6	16.6	16.0
	SD	2.5	3.5	2.9	2.0	3.2	1.6	6.7	2.6	2.5	4.9
	N	150	150	150	150	150	150	150	150	150	150
15	M	24.6	15.9	18.4	20.1	24.7	15.2	47.7	13.9	16.8	16.6
	SD	2.5	3.8	2.8	2.3	3.2	1.7	8.1	2.8	2.8	4.3
	N	150	150	150	150	150	150	150	150	150	150
16	M	25.4	17.7	20.0	21.2	26.6	15.6	51.1	14.9	17.9	18.8
	SD	2.4	3.9	3.1	2.9	3.3	1.8	6.4	3.4	2.8	4.3
	N	150	150	150	150	150	150	150	150	150	150

Table

Mean and SD of Raw Scores- All students of Barisal Division

		Non Verbal					Verbal				
Age		RS_NV_FR	RS_NV_KN	RS_NV_QR	RS_NV_VS	RS_NV_WM	RS_V_FR	RS_V_KN	RS_V_QR	RS_V_VS	RS_V_WM
6	M	11.2	11.1	14.7	13.2	18.5	10.3	27.4	11.0	13.7	11.4
	SD	1.4	1.8	2.1	4.0	4.6	2.8	6.3	2.1	2.3	4.8
	N	50	50	50	50	50	50	50	50	50	50
7	M	15.4	11.6	14.3	16.3	18.1	10.1	26.9	11.2	13.2	8.8
	SD	1.4	1.9	1.7	2.6	3.2	2.7	7.3	2.5	2.7	4.5
	N	50	50	50	50	50	50	50	50	50	50
8	M	15.9	12.9	15.6	17.9	20.6	11.6	32.7	11.8	13.9	13.2
	SD	1.2	2.0	1.7	2.4	2.9	1.8	5.0	0.9	2.6	4.0
	N	50	50	50	50	50	50	50	50	50	50
9	M	15.5	12.4	15.2	16.9	19.4	11.3	31.7	11.9	14.5	10.8
	SD	1.0	1.9	1.4	2.6	2.3	1.8	5.4	1.7	2.2	4.4
	N	50	50	50	50	50	50	50	50	50	50
10	M	15.9	13.0	15.2	18.0	20.5	13.4	42.4	12.5	15.3	11.5
	SD	0.9	2.0	1.1	1.5	3.3	1.1	5.4	0.8	1.6	6.3
	N	50	50	50	50	50	50	50	50	50	50
11	M	16.0	12.7	15.4	17.5	21.0	13.8	41.8	12.9	15.3	12.6
	SD	1.0	2.4	1.7	2.3	3.8	1.5	7.1	1.4	1.2	6.3
	N	50	50	50	50	50	50	50	50	50	50
12	M	16.0	12.8	15.5	18.1	19.7	13.5	43.7	12.2	15.4	11.9
	SD	1.0	1.8	0.9	1.2	2.5	0.9	5.8	0.6	1.6	4.8
	N	50	50	50	50	50	50	50	50	50	50
13	M	21.4	13.9	16.1	18.2	20.9	13.2	45.1	12.3	15.5	12.2
	SD	2.7	2.7	1.8	2.0	2.8	1.3	5.8	0.9	1.4	4.3
	N	50	50	50	50	50	50	50	50	50	50
14	M	22.2	14.0	17.0	18.6	22.1	14.1	45.9	13.1	16.1	12.9
	SD	2.4	2.6	1.9	1.9	2.8	1.3	5.5	2.1	2.4	5.4
	N	50	50	50	50	50	50	50	50	50	50
15	M	24.0	16.5	18.7	19.9	24.8	15.6	48.6	15.5	18.6	17.6
	SD	2.4	4.0	3.1	2.3	3.7	1.7	7.8	2.9	2.8	4.9
	N	50	50	50	50	50	50	50	50	50	50
16	M	25.2	18.1	21.4	21.9	28.8	15.8	51.9	17.0	18.9	22.2
	SD	2.6	4.4	3.3	2.8	2.6	2.4	7.3	3.6	3.2	3.5
	N	50	50	50	50	50	50	50	50	50	50

Table

Mean and SD of Raw Scores- All Students of Chittagong Division

		Non Verbal					Verbal				
Age		RS_NV_FR	RS_NV_KN	RS_NV_QR	RS_NV_VS	RS_NV_WM	RS_V_FR	RS_V_KN	RS_V_QR	RS_V_VS	RS_V_WM
6	M	11.9	11.4	15.3	15.4	20.7	11.0	30.1	11.8	13.4	12.4
	SD	1.3	1.7	2.0	1.8	3.5	1.7	4.9	0.8	2.1	3.4
	N	50	50	50	50	50	50	50	50	50	50
7	M	16.5	12.2	14.5	17.5	19.2	11.3	27.9	11.6	11.9	9.3
	SD	0.6	1.1	0.7	0.7	1.2	1.2	5.6	0.6	0.7	1.0
	N	50	50	50	50	50	50	50	50	50	50
8	M	16.6	12.6	14.5	17.8	19.2	11.7	29.1	11.8	12.5	9.7
	SD	0.6	1.3	0.8	1.0	1.2	1.3	6.0	0.4	1.2	2.1
	N	50	50	50	50	50	50	50	50	50	50
9	M	16.6	13.2	14.9	18.1	20.1	12.4	31.6	11.9	13.2	11.0
	SD	0.6	1.5	1.1	0.9	1.8	1.6	5.8	0.4	2.1	3.7
	N	50	50	50	50	50	50	50	50	50	50
10	M	16.6	12.8	15.0	18.2	19.8	13.7	41.6	12.0	14.3	12.2
	SD	0.6	1.3	1.1	1.0	1.7	1.5	6.8	0.9	1.9	2.3
	N	50	50	50	50	50	50	50	50	50	50
11	M	16.7	13.3	15.3	18.5	20.8	14.5	43.4	12.1	15.5	12.7
	SD	0.6	1.9	1.1	0.9	1.3	1.1	7.3	0.7	2.0	4.2
	N	50	50	50	50	50	50	50	50	50	50
12	M	16.8	13.8	15.4	18.8	20.7	14.3	45.7	12.2	15.2	14.1
	SD	0.4	1.7	1.2	1.3	1.8	0.6	6.5	0.7	1.4	3.8
	N	50	50	50	50	50	50	50	50	50	50
13	M	25.0	15.9	18.0	19.9	24.0	15.0	49.5	12.9	16.1	15.3
	SD	2.1	3.6	2.2	2.2	2.5	1.3	2.9	2.2	2.5	4.0
	N	50	50	50	50	50	50	50	50	50	50
14	M	25.6	17.3	19.0	20.4	25.1	15.3	49.8	13.2	16.9	15.4
	SD	2.5	2.9	2.0	2.4	2.8	1.6	2.7	2.7	2.5	4.3
	N	50	50	50	50	50	50	50	50	50	50
15	M	26.3	16.8	19.3	21.1	25.2	15.1	49.7	12.7	16.5	14.7
	SD	1.8	3.7	2.5	2.3	2.4	1.1	5.4	1.0	1.7	2.0
	N	50	50	50	50	50	50	50	50	50	50
16	M	26.8	20.1	21.1	22.5	27.6	16.6	54.2	16.5	19.0	19.7
	SD	1.7	3.5	3.7	3.7	3.0	2.1	8.1	4.8	3.4	4.7
	N	50	50	50	50	50	50	50	50	50	50

Table

Mean and SD of Raw Scores- All students of Dhaka Division

		Non Verbal					Verbal				
Age		**RS_NV_FR**	**RS_NV_KN**	**RS_NV_QR**	**RS_NV_VS**	**RS_NV_WM**	**RS_V_FR**	**RS_V_KN**	**RS_V_QR**	**RS_V_VS**	**RS_V_WM**
6	M	11.4	10.6	14.5	14.2	19.1	9.7	26.9	11.5	12.9	10.7
	SD	1.5	1.9	2.1	2.8	3.3	2.3	6.0	1.4	2.0	4.4
	N	50	50	50	50	50	50	50	50	50	50
7	M	15.1	11.8	14.6	16.7	19.0	9.9	30.1	11.4	13.7	11.5
	SD	1.9	2.0	1.6	2.1	3.0	3.4	5.6	1.0	2.1	4.6
	N	50	50	50	50	50	50	50	50	50	50
8	M	15.7	12.1	14.9	17.4	19.8	11.3	29.8	11.7	14.1	11.2
	SD	1.3	1.9	1.8	2.3	2.9	2.4	6.5	1.3	2.4	3.7
	N	50	50	50	50	50	50	50	50	50	50
9	M	16.0	13.4	15.7	17.8	20.5	12.4	32.3	12.1	15.0	12.6
	SD	1.0	1.9	1.6	1.3	2.1	2.0	7.1	0.8	1.7	3.4
	N	50	50	50	50	50	50	50	50	50	50
10	M	16.1	12.9	15.5	17.9	20.9	13.4	39.8	12.1	15.1	12.5
	SD	0.9	1.7	1.0	1.4	2.2	1.3	8.7	0.4	1.7	2.8
	N	50	50	50	50	50	50	50	50	50	50
11	M	16.3	14.1	16.0	18.6	22.5	14.1	44.2	12.4	15.4	15.2
	SD	1.4	2.8	1.6	1.4	2.7	1.7	6.6	1.5	2.0	4.4
	N	50	50	50	50	50	50	50	50	50	50
12	M	16.2	13.3	16.1	18.4	21.6	13.8	44.9	12.4	15.6	13.8
	SD	0.8	1.7	1.1	1.9	2.3	1.1	6.8	0.6	1.4	2.6
	N	50	50	50	50	50	50	50	50	50	50
13	M	23.1	14.6	17.3	19.0	23.2	14.1	47.6	12.5	16.5	14.9
	SD	2.5	2.5	2.3	1.2	2.6	1.5	4.0	1.1	2.1	3.1
	N	50	50	50	50	50	50	50	50	50	50
14	M	24.6	15.2	18.0	19.6	24.8	14.4	48.3	13.1	16.9	17.1
	SD	1.9	3.7	2.0	1.6	2.7	1.3	6.7	1.9	2.2	2.7
	N	50	50	50	50	50	50	50	50	50	50
15	M	24.6	16.2	18.8	19.9	25.5	15.2	50.3	13.8	17.7	18.0
	SD	2.2	2.9	2.6	1.4	2.6	1.8	6.6	2.6	2.5	3.0
	N	50	50	50	50	50	50	50	50	50	50
16	M	24.7	16.0	18.9	20.0	25.2	14.4	49.2	13.4	17.4	17.1
	SD	2.7	3.3	2.7	1.9	2.8	1.8	6.0	2.1	2.7	2.8
	N	50	50	50	50	50	50	50	50	50	50

Table

Mean and SD of Raw Scores- All Students of Khulna Division

		Non Verbal					Verbal				
Age		**RS_NV_FR**	**RS_NV_KN**	**RS_NV_QR**	**RS_NV_VS**	**RS_NV_WM**	**RS_V_FR**	**RS_V_KN**	**RS_V_QR**	**RS_V_VS**	**RS_V_WM**
6	**M**	11.2	10.6	14.1	13.0	16.5	9.1	25.7	12.0	13.1	12.3
	SD	1.9	1.9	1.9	3.2	5.1	2.8	6.5	2.2	2.5	2.9
	N	50	50	50	50	50	50	50	50	50	50
7	**M**	15.7	12.2	15.4	16.6	19.3	10.9	31.3	12.5	14.9	15.2
	SD	1.9	1.9	1.9	3.2	4.0	3.8	7.6	1.2	2.2	3.6
	N	50	50	50	50	50	50	50	50	50	50
8	**M**	16.0	11.9	15.2	16.9	19.3	9.8	32.3	12.2	14.8	14.8
	SD	1.1	2.2	2.1	2.7	3.8	3.9	4.7	1.5	2.6	4.4
	N	50	50	50	50	50	50	50	50	50	50
9	**M**	15.8	13.5	15.8	15.8	18.6	9.8	30.5	11.6	13.8	14.7
	SD	1.2	2.4	2.4	5.0	4.5	3.8	6.3	2.7	2.9	5.1
	N	50	50	50	50	50	50	50	50	50	50
10	**M**	16.3	13.3	16.0	17.9	19.7	12.7	42.2	12.6	15.3	16.0
	SD	0.9	2.0	2.0	1.6	3.7	1.4	8.4	1.2	1.7	4.0
	N	50	50	50	50	50	50	50	50	50	50
11	**M**	16.4	13.7	15.8	17.9	19.1	13.1	43.0	13.1	15.9	14.5
	SD	0.9	2.6	2.0	1.7	2.6	2.7	6.8	1.6	1.6	4.5
	N	50	50	50	50	50	50	50	50	50	50
12	**M**	16.6	14.6	17.0	18.4	22.7	14.6	45.6	13.8	16.3	14.4
	SD	0.6	3.5	2.1	1.4	4.1	1.7	6.4	2.4	1.3	3.6
	N	50	50	50	50	50	50	50	50	50	50
13	**M**	24.9	16.8	18.2	19.6	25.1	15.1	48.6	14.5	16.9	16.9
	SD	2.2	2.8	2.9	2.7	2.7	1.6	4.2	2.3	2.5	4.9
	N	50	50	50	50	50	50	50	50	50	50
14	**M**	25.9	18.9	20.5	21.1	26.3	15.9	51.4	16.5	19.0	20.7
	SD	2.2	3.3	2.6	2.0	2.9	1.9	6.1	3.0	2.6	6.1
	N	50	50	50	50	50	50	50	50	50	50
15	**M**	26.7	19.4	21.0	22.6	27.9	17.1	53.4	17.9	19.6	22.1
	SD	1.9	5.1	4.0	3.6	3.3	2.1	8.6	4.0	3.2	4.1
	N	50	50	50	50	50	50	50	50	50	50
16	**M**	27.0	20.3	22.4	22.8	28.7	17.3	54.5	18.1	20.1	22.8
	SD	1.9	3.4	2.7	2.5	2.8	3.4	9.2	4.0	2.8	4.2
	N	50	50	50	50	50	50	50	50	50	50

Table

Mean and SD of Raw Scores- All Students of Rajshahi Division

		Non Verbal					Verbal				
Age		RS_NV_FR	RS_NV_KN	RS_NV_QR	RS_NV_VS	RS_NV_WM	RS_V_FR	RS_V_KN	RS_V_QR	RS_V_VS	RS_V_WM
6	M	11.8	10.2	14.0	14.7	18.1	11.0	27.4	11.4	11.2	7.7
	SD	1.0	1.0	0.7	2.4	1.2	1.9	4.0	1.3	1.4	1.3
	N	50	50	50	50	50	50	50	50	50	50
7	M	15.2	12.5	13.8	17.0	18.3	10.2	27.5	11.2	11.2	8.5
	SD	1.5	1.5	0.8	1.4	1.5	2.3	5.5	1.0	1.9	2.1
	N	50	50	50	50	50	50	50	50	50	50
8	M	16.0	11.8	14.0	16.7	18.5	11.1	29.0	11.7	12.2	8.5
	SD	0.7	1.1	0.9	1.3	1.3	1.1	3.8	0.5	1.4	1.3
	N	50	50	50	50	50	50	50	50	50	50
9	M	15.8	13.0	14.7	17.3	19.5	11.6	30.2	11.8	13.0	10.3
	SD	0.8	1.4	0.9	1.3	1.6	1.4	5.7	0.4	1.3	2.9
	N	50	50	50	50	50	50	50	50	50	50
10	M	16.2	12.8	14.5	17.6	19.7	13.7	40.0	11.9	13.2	11.8
	SD	0.7	1.7	0.7	1.0	1.5	0.9	5.0	0.4	1.0	1.8
	N	50	50	50	50	50	50	50	50	50	50
11	M	16.2	13.1	14.9	17.9	20.9	14.0	40.4	12.1	13.7	14.5
	SD	1.6	2.7	1.4	1.6	2.9	1.3	7.5	0.5	1.9	4.1
	N	50	50	50	50	50	50	50	50	50	50
12	M	16.5	12.4	14.8	18.1	20.1	13.9	42.0	12.2	13.8	13.4
	SD	2.0	1.9	1.3	1.1	2.0	0.9	6.9	0.5	1.4	2.9
	N	50	50	50	50	50	50	50	50	50	50
13	M	23.8	14.7	16.3	18.9	23.1	14.3	45.8	12.4	14.5	14.4
	SD	1.9	2.8	1.4	1.2	2.3	0.8	4.1	1.1	1.6	2.0
	N	50	50	50	50	50	50	50	50	50	50
14	M	24.0	15.6	17.4	19.4	24.0	14.7	46.5	13.0	15.3	15.5
	SD	2.7	3.0	2.6	2.4	3.1	1.2	6.8	1.4	1.8	2.6
	N	50	50	50	50	50	50	50	50	50	50
15	M	24.6	15.4	17.4	19.8	24.1	14.6	46.9	13.2	15.2	16.1
	SD	2.4	4.0	2.7	2.2	3.2	1.4	6.6	2.0	1.8	2.9
	N	50	50	50	50	50	50	50	50	50	50
16	M	25.2	17.3	19.1	20.9	25.8	15.4	49.9	14.3	16.3	17.8
	SD	1.8	3.4	2.7	2.4	3.1	1.5	5.2	2.8	2.1	3.5
	N	50	50	50	50	50	50	50	50	50	50

Table

Mean and SD of Raw Scores- All Students of Sylhet Division

		Non Verbal					Verbal				
Age		RS_NV_FR	RS_NV_KN	RS_NV_QR	RS_NV_VS	RS_NV_WM	RS_V_FR	RS_V_KN	RS_V_QR	RS_V_VS	RS_V_WM
6	M	11.6	9.9	13.7	14.2	17.8	10.5	26.6	11.5	12.1	8.7
	SD	1.0	0.9	0.8	2.5	1.5	1.5	4.1	0.7	1.1	2.0
	N	50	50	50	50	50	50	50	50	50	50
7	M	15.0	11.7	13.9	16.7	18.6	11.0	29.5	11.5	12.7	9.9
	SD	1.2	1.3	1.0	1.4	1.7	1.8	4.0	0.9	1.6	2.5
	N	50	50	50	50	50	50	50	50	50	50
8	M	15.9	12.0	14.2	17.4	19.1	11.2	29.8	11.6	13.1	11.5
	SD	0.9	1.6	1.6	1.7	2.4	1.8	5.6	1.3	2.5	4.2
	N	50	50	50	50	50	50	50	50	50	50
9	M	15.6	11.8	14.2	17.2	19.0	11.6	28.3	11.8	12.5	9.8
	SD	0.9	1.3	0.9	1.4	1.4	1.4	6.3	0.4	1.0	1.5
	N	50	50	50	50	50	50	50	50	50	50
10	M	15.7	12.6	14.4	17.7	19.9	13.3	39.4	11.7	13.3	10.0
	SD	1.1	1.4	1.0	1.4	2.1	1.0	4.9	1.3	1.5	3.1
	N	50	50	50	50	50	50	50	50	50	50
11	M	16.0	13.1	14.9	18.1	20.5	13.9	42.2	12.0	13.7	10.9
	SD	1.1	1.7	0.9	1.3	2.0	1.0	4.8	0.2	1.5	2.8
	N	50	50	50	50	50	50	50	50	50	50
12	M	16.4	13.0	15.4	18.4	20.8	13.9	40.7	12.0	14.1	10.9
	SD	1.9	2.1	1.4	2.4	2.5	0.9	7.9	0.5	1.7	4.2
	N	50	50	50	50	50	50	50	50	50	50
13	M	22.4	13.6	15.9	18.7	21.4	14.0	44.7	11.9	14.4	11.5
	SD	2.9	2.1	1.4	1.3	2.2	0.9	6.8	0.5	1.4	3.0
	N	50	50	50	50	50	50	50	50	50	50
14	M	23.4	14.9	16.7	19.3	23.1	14.3	45.6	12.4	15.0	13.0
	SD	2.8	3.0	2.7	1.4	3.0	1.3	5.9	1.2	1.9	4.2
	N	50	50	50	50	50	50	50	50	50	50
15	M	23.2	15.1	17.3	19.6	23.0	14.7	45.1	12.3	15.3	13.5
	SD	2.9	2.5	2.3	1.3	2.7	1.1	8.3	0.9	1.9	3.1
	N	50	50	50	50	50	50	50	50	50	50
16	M	24.1	16.1	18.4	19.8	24.2	15.2	48.5	12.8	16.7	15.5
	SD	3.0	3.3	2.8	1.6	3.4	1.4	7.5	1.9	2.0	3.1
	N	50	50	50	50	50	50	50	50	50	50

Table

Weighted Mean and Weighted SD of Raw Scores- All Students of Barisal Division

	Non verbal										Verbal									
Age	RS_NV_FR		RS_NV_KN		RS_NV_QR		RS_NV_VS		RS_NV_WM		RS_V_FR		RS_V_KN		RS_V_QR		RS_V_VS		RS_V_WM	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
6	11.3	1.3	11.4	1.7	15.0	1.9	14.4	3.2	19.6	3.4	11.1	2.1	28.9	5.5	11.4	1.4	14.1	2.2	13.4	4.9
7	15.6	1.3	11.9	1.7	14.5	1.5	16.7	2.2	18.7	2.5	10.8	2.7	28.8	6.1	11.8	2.9	13.7	2.5	11.1	4.9
8	15.9	1.1	13.3	2.0	15.7	1.8	18.2	2.4	21.0	2.7	11.9	1.7	33.5	4.4	11.9	0.9	14.4	2.5	14.3	3.8
9	15.6	1.0	12.7	2.0	15.3	1.4	17.3	2.1	19.7	2.2	11.6	1.7	32.6	5.3	12.1	1.5	14.8	1.9	12.6	4.7
10	16.0	0.9	13.3	2.1	15.3	1.1	18.2	1.5	21.0	3.5	13.5	1.1	43.1	4.9	12.5	0.8	15.5	1.5	14.9	6.5
11	16.0	1.0	13.1	2.3	15.6	1.6	17.8	1.9	21.7	3.6	13.9	1.5	43.0	6.0	13.0	1.5	15.4	1.2	15.7	6.4
12	16.1	0.9	13.0	1.9	15.6	0.9	18.1	1.2	20.0	2.7	13.5	0.9	44.5	5.3	12.2	0.6	15.6	1.5	13.8	4.4
13	21.8	2.9	14.4	2.8	16.3	1.9	18.4	1.8	21.3	2.9	13.3	1.3	45.8	5.2	12.3	1.0	15.6	1.4	13.7	3.8
14	22.5	2.4	14.4	2.7	17.2	2.0	18.8	2.0	22.5	2.8	14.2	1.3	46.5	5.9	13.5	2.4	16.4	2.3	15.1	4.7
15	24.2	2.3	17.4	3.8	19.2	3.1	20.2	2.3	25.4	3.4	15.8	1.7	49.8	6.3	16.1	2.9	19.0	2.9	19.0	4.2
16	25.4	2.5	19.2	4.3	21.9	3.1	22.2	2.8	29.0	2.6	16.2	2.4	52.9	6.8	17.8	3.8	19.4	3.1	22.8	3.4

Table

Weighted Mean and Weighted SD of Raw Scores- All students of Chittagong Division

	Non verbal										Verbal									
Age	RS_NV_FR		RS_NV_KN		RS_NV_QR		RS_NV_VS		RS_NV_WM		RS_V_FR		RS_V_KN		RS_V_QR		RS_V_VS		RS_V_WM	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
6	12.0	1.3	11.7	1.8	15.5	2.1	15.7	1.9	21.3	3.5	11.2	1.6	30.8	4.3	11.9	0.9	13.7	2.1	13.3	3.3
7	16.5	0.6	12.3	1.1	14.5	0.7	17.5	0.7	19.3	1.2	11.5	1.2	29.0	4.4	11.6	0.6	12.0	0.7	9.4	1.0
8	16.7	0.6	12.8	1.3	14.6	0.8	17.9	1.0	19.3	1.2	11.8	1.3	30.3	4.8	11.8	0.4	12.6	1.3	10.1	2.1
9	16.6	0.6	13.4	1.5	15.0	1.1	18.2	0.9	20.2	1.8	12.6	1.5	32.6	4.8	11.9	0.3	13.5	2.2	12.2	4.3
10	16.6	0.6	13.0	1.2	15.1	1.1	18.3	1.0	19.9	1.6	13.9	1.4	42.7	5.5	12.1	1.1	14.5	1.8	12.6	2.5
11	16.7	0.5	13.6	1.6	15.4	1.1	18.5	0.9	20.9	1.3	14.5	1.1	44.6	5.9	12.1	0.7	15.7	2.1	14.1	5.0
12	16.9	0.4	14.0	1.7	15.5	1.2	18.9	1.4	20.9	1.8	14.3	0.6	46.6	5.0	12.2	0.8	15.3	1.3	15.1	4.2
13	25.1	2.0	16.7	3.0	18.2	2.2	20.1	2.2	24.2	2.5	15.1	1.3	49.6	2.9	13.3	2.5	16.5	2.6	16.4	4.6
14	25.8	2.1	17.8	2.6	19.2	1.9	20.7	2.1	25.5	2.8	15.4	1.6	50.0	2.7	13.8	3.2	17.3	26	16.6	4.9
15	26.5	1.7	17.6	3.1	19.6	2.4	21.3	2.3	25.4	2.3	15.2	1.1	50.3	5.2	12.8	1.0	16.7	17	15.0	2.0
16	26.9	1.6	20.7	3.1	21.8	3.1	23.1	3.2	27.9	2.9	16.9	2.2	55.4	7.9	17.8	5.2	19.6	3.5	20.8	4.6

Table

Weighted Mean and Weighted SD of Raw Scores- All students of Dhaka Division

	Non verbal										Verbal									
Age	RS_NV_FR		RS_NV_KN		RS_NV_QR		RS_NV_VS		RS_NV_WM		RS_V_FR		RS_V_KN		RS_V_QR		RS_V_VS		RS_V_WM	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
6	11.5	1.9	11.0	1.8	14.3	1.9	13.8	2.7	18.0	4.9	9.9	1.8	27.3	5.7	12.4	2.0	13.6	2.2	13.0	2.8
7	15.9	1.6	12.5	1.7	15.6	2.0	17.2	2.5	20.1	3.8	12.1	2.9	33.1	6.0	12.6	1.3	15.3	2.2	16.0	3.5
8	16.1	1.1	12.3	1.7	15.5	1.9	17.3	2.1	20.0	3.7	11.3	2.9	32.9	4.3	12.3	1.6	15.3	2.8	16.1	4.2
9	15.9	1.2	13.9	2.6	16.1	2.7	17.4	4.0	19.6	3.7	11.2	2.9	31.7	5.5	12.2	2.4	14.4	2.9	16.5	4.7
10	16.4	0.9	13.6	1.8	16.3	2.0	18.0	1.5	20.4	3.9	12.9	1.4	43.8	6.6	12.7	1.3	15.5	1.7	16.9	3.8
11	16.5	0.8	14.2	2.3	16.0	2.0	18.1	1.6	19.4	2.6	13.7	3.2	44.0	5.8	13.3	1.7	16.1	1.6	15.9	4.4
12	16.6	0.6	15.4	3.4	17.2	2.2	18.5	1.4	23.4	4.3	14.8	1.8	46.5	5.3	14.3	2.6	16.4	1.3	15.3	3.4
13	25.1	2.0	17.3	2.6	18.6	2.5	20.0	1.8	25.4	2.7	15.2	1.6	49.0	4.1	14.8	2.3	17.3	2.6	18.3	4.7
14	26.0	2.0	19.4	3.2	20.8	2.5	21.3	1.9	26.6	2.7	16.1	1.9	52.1	5.5	17.1	3.0	19.3	2.5	22.5	56
15	26.9	19	20.7	4.0	21.8	3.6	23.1	3.4	28.3	3.1	17.4	2.0	54.7	8.3	18.8	4.0	20.1	3.1	22.8	3.5
16	27.1	1.8	20.9	3.3	22.7	2.7	23.1	2.5	28.9	2.7	17.9	3.6	56.0	8.1	19.0	3.9	20.5	2.8	23.6	3.8

Table

Weighted Mean and Weighted SD of Raw Scores- All Students of Khulna Division

	Non verbal										Verbal									
Age	RS_NV_FR		RS_NV_KN		RS_NV_QR		RS_NV_VS		RS_NV_WM		RS_V_FR		RS_V_KN		RS_V_QR		RS_V_VS		RS_V_WM	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
6	11.6	1.5	11.0	1.9	14.8	2.1	14.7	2.2	19.6	3.4	10.3	1.8	28	5.2	1.6	1.4	13.2	1.9	12.4	4.9
7	15.4	1.6	12.1	1.9	14.7	1.7	17.0	1.9	19.5	3.0	11.1	2.3	31.1	4.7	1.5	0.8	14.0	1.8	13.3	4.4
8	15.8	1.2	12.4	1.8	15.1	1.9	17.7	2.0	20.2	2.8	11.8	2.0	31.2	5.5	1.8	1.3	14.6	2.2	12.3	3.7
9	16.1	0.9	13.6	1.9	15.9	1.7	17.9	1.3	20.7	2.2	12.7	1.9	33.9	5.7	2.2	0.8	15.2	1.7	13.5	3.2
10	16.2	0.8	13.1	1.7	15.5	1.0	18.0	1.3	21.2	2.2	13.5	1.4	41.6	7.2	2.1	0.4	15.2	1.6	13.1	2.3
11	16.5	1.5	14.6	3.1	16.2	1.6	18.7	1.4	22.8	2.7	14.3	1.8	45.0	5.8	2.6	1.9	15.6	2.0	16.5	4.5
12	16.3	0.8	13.7	1.8	16.2	1.1	18.6	1.6	21.9	2.4	14.1	1.2	45.7	6.3	2.5	0.9	15.7	1.4	14.7	2.4
13	23.4	2.3	14.8	2.0	17.5	2.4	19.1	1.2	23.6	2.5	14.2	1.5	48.2	3.9	2.5	1.1	16.7	2.2	15.6	2.7
14	24.6	2.0	16.4	3.3	18.3	2.1	19.8	1.6	25.0	2.6	14.5	1.3	49.5	5.5	3.4	2.1	17.2	2.1	17.3	2.6
15	24.5	2.1	16.2	2.7	18.8	2.5	19.9	1.4	25.3	2.5	15.3	1.9	50.4	5.5	4.2	2.9	17.8	2.6	18.1	3.2
16	25.0	2.4	16.7	3.5	19.3	2.6	20.2	1.7	25.5	2.7	14.7	1.8	50.0	5.2	3.7	2.3	17.8	2.7	17.6	2.9

Table

Weighted Mean and Weighted SD of Raw Scores- All Students of Rajshahi Division

	Non verbal										Verbal									
Age	RS_NV_FR		RS_NV_KN		RS_NV_QR		RS_N_VS		RS_NV_WM		RS_V_FR		RS_V_KN		RS_V_QR		RS_V_VS		RS_V_WM	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
6	11.8	1.0	10.3	1.0	14.0	0.7	15.1	1.6	18.2	1.2	11.1	2.1	27.9	3.1	11.6	1.5	11.4	1.5	7.9	1.5
7	15.3	1.3	12.6	1.4	13.9	0.8	17.1	1.3	18.5	1.4	10.8	2.7	28.5	4.2	11.3	0.9	11.6	1.9	9.0	2.4
8	16.0	0.7	12.0	1.0	14.1	0.9	16.9	1.0	18.7	1.3	11.9	1.7	29.6	3.2	11.7	0.5	12.4	1.2	8.6	1.4
9	15.8	0.8	13.1	1.4	14.7	0.9	17.4	1.3	19.7	1.6	11.6	1.7	31.2	4.6	11.9	0.4	13.1	1.2	11.1	3.2
10	16.2	0.7	12.7	1.5	14.5	0.8	17.7	1.0	19.7	1.5	13.5	1.1	40.6	4.5	11.9	0.4	13.4	0.9	11.9	1.7
11	16.4	1.9	13.7	2.8	15.1	1.4	18.0	1.5	21.3	3.0	13.9	1.5	41.8	6.0	12.2	0.6	14.0	1.9	15.6	4.5
12	16.8	2.4	12.8	2.1	15.0	1.4	18.2	1.2	20.5	1.9	13.5	0.9	43.5	6.0	12.2	0.5	14.0	1.4	14.1	3.5
13	23.9	1.7	15.2	2.6	16.4	1.4	19.0	1.2	23.3	2.2	13.3	1.3	46.2	3.6	12.5	1.2	14.7	1.6	14.7	2.0
14	24.2	2.5	15.9	2.8	17.5	2.6	19.4	2.2	24.2	3.0	14.2	1.3	47.1	5.6	13.2	1.5	15.5	1.7	15.9	2.5
15	24.8	2.2	16.5	3.6	18.0	2.7	20.0	2.1	24.7	3.1	15.8	1.7	47.7	5.4	13.4	2.2	15.4	2.0	16.7	3.0
16	25.3	1.7	18.0	3.0	19.5	2.6	21.1	2.5	26.2	2.9	16.2	2.4	50.4	4.9	14.9	3.0	16.6	2.2	18.5	3.5

Table

Weighted Mean and Weighted SD of Raw Scores- All students of Sylhet Division

	Non verbal										Verbal									
Age	RS_NV_FR		RS_NV_KN		RS_NV_QR		RS_NV_VS		RS_NV_WM		RS_V_FR		RS_V_KN		RS_V_QR		RS_V_VS		RS_V_WM	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
6	11.6	1.0	10.0	0.9	137	0.8	14.6	2.0	17.9	1.5	10.7	1.4	27.2	3.4	11.5	0.7	12.2	1.1	9.2	2.2
7	15.1	1.1	11.8	1.3	14.0	0.9	16.8	1.4	18.8	1.7	11.3	1.4	30.0	3.3	11.6	0.7	12.9	1.6	10.6	2.8
8	15.9	0.9	12.2	1.6	14.4	1.5	17.5	1.5	19.4	2.3	11.5	1.5	30.8	4.8	11.7	1.0	13.6	2.5	13.0	4.3
9	15.6	0.9	12.0	1.4	14.2	1.0	17.3	1.4	19.0	1.5	11.7	1.4	29.7	5.1	11.9	0.4	12.6	1.0	10.0	1.5
10	15.8	1.3	12.4	1.3	14.5	0.8	17.8	1.3	20.0	2.0	13.4	0.9	40.3	3.8	11.8	0.7	13.4	1.5	11.0	2.7
11	16.0	1.0	13.2	1.6	14.9	0.9	18.2	1.2	20.8	2.0	14.0	1.0	42.7	4.3	12.0	0.2	14.0	1.5	11.8	3.1
12	16.3	1.7	13.2	2.1	15.6	1.4	18.7	1.9	20.9	2.5	13.9	0.8	42.2	6.9	11.9	0.4	14.1	1.6	12.0	4.2
13	22.7	3.1	13.9	2.2	16.0	1.5	18.8	1.3	21.7	2.3	13.9	1.0	45.6	6.0	11.9	0.5	14.5	1.4	12.2	3.0
14	23.7	2.5	15.5	2.6	17.2	2.0	19.3	1.4	23.2	2.7	14.5	1.2	46.6	4.9	12.4	1.0	15.3	1.7	14.4	4.0
15	23.5	2.8	15.3	2.3	17.6	2.2	19.6	1.3	23.3	2.6	14.7	1.1	46.5	6.6	12.3	0.9	15.6	1.8	14.2	3.0
16	24.4	2.9	16.8	3.3	18.9	2.8	19.9	1.6	24.6	3.3	15.4	1.4	49.6	6.0	13.1	2.2	17.0	2.0	16.1	3.1

Table

Scaled Score Equivalents of Raw Score (Non verbal) for 6 Years Children

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
1	0-6	0-4	0-6	0-2	0-5
2	7-9	5-8	7-13	3-11	6-15
3	10	9		12	16
4	11	-		13	17
5	-	10		14	
6	-		14		18
7	-				
8	-				
9	12			15	
10					19
11					
12					20
13					
14					21
15	13	12	16	16	22
16		13	17	17	23-24
17		14	18	18	25-26
18	15-16	15-17	19	18-20	27-30
19	17-up	18-up	20-up	21-up	31- above

Table

Scaled Score Equivalents of Raw Score (Verbal) for 6 Years Children

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
1	0-1	0-11	0-2	0-4	0-4
2	2-8	12 to 14	3-10	5-10	5-7
3		20 to 23	11		8
4	9	24		11	
5	10	28		12	
6		29			
7		30			9
8	11				10
9		31	12	13	11
10					12
11		32		14	13
12				15	14
13					15

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
14	12				16
15		33 to 34			17
16		35	13		
17	13	36 to 38	14	16	19
18	14-16	38 to 39	15-18	17-21	20-24
19	17- up	39-above	19-up	22-up	25-up

Table

Scaled Score Equivalents of Raw Score (Non verbal) for 7 Years Children

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
1	0-8	0-4	0-4	0-7	0-5
2	9-13	5	5	8	6-14
3	14	6	6	9-12	15
4	15	7-10	7	13	16
5		11	8-13	14	
6		12	14	15	
7	16			16	
8					17
9					
10		13			
11		14	15	17	
12		15		18	18
13	17				
14			16		
15			17		
16			18		19
17		15-16	19-20	19-20	20-21
18		17-23	21	20-24	22
19	31-36	24-30	22-30	25-34	25-34

Table

Scaled Score Equivalents of Raw Score (Verbal) for 7 Years Children

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
1	0-3	0 to 11	0-4	0-7	0-1
2	4-7	12 to 14	5-10	7-11	1-7

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
3	8	15 to 17	11	12	8
4	9	18 to 23	12	13	9
5	10	24 to 26	13	14	
6	11	27			
7	12	28			
8	13	29			10
9	14	30			11
10	15	31			12
11	16	32		14	13
12	17	33			14
13	18	34			15-16
14	19	35			17
15	20	36		15	18-19
16	21	37		16	20
17	22-24				21-22
18	25-28	38	14-24	19-21	
19	29-30	39 above	25-30	21-34	23-34

Table

Scaled Score Equivalents of Raw Score (Non verbal) for 8 Years Children

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
1	0-6	0-7	0 to 6	0 to 5	0 to 7
2	7-11	8	7 to 13	6 to 14	8 to 14
3	12	9		15	15
4	13	10	14	16	16
5	14				17
6				17	18
7		11			19
8		12-13			20
9		14	15	18	21
10		15-16			
11		17		19	22
12	15	18-19			23
13	16	20	16		
14	17			20	
15	18		27		24
16					25
17					26
18	19-21	21-22	18 to 21	21 to 26	26
19	21-36	22-30	21 above	26 above	26 above

Table

Scaled Score Equivalents of Raw Score (Verbal) for 8 Years Children

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
1	0 to 1	0 to 12	0-5	0-6	0-8
2	2 to 9	13 to 19	6-11	6-11	
3	10	20		12	
4	11	21			9
5	12	22	12	13	10
6	13	23 to 28		14	
7	14	29			
8	15	30	13		11
9	16	31			12-13
10	17	32			14
11	18	33		15	15
12	19	34		16	16
13		35		17	17
14		36		18	
15		37			
16		38			
17		39	14-18		
18		40 to 41		19-23	19-23
19		41 above	18-30	23-34	24-334

Table

Scaled Score Equivalents of Raw Score (Non verbal) for 9 Years Children

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
1	0-12	0-7	0-9	0-3	0-2
2	13	8-10	10	5	4-6
3	14	11	11	6	7-10
4		12	12	7	10-11
5	15		13	8	11-18
6	16		14	9	19
7		13	15	10	20
8		14	16	11	21
9		15		12	
10				13	
11				14	
12	17			15-18	
13				19	

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
14				20	22
15			17	21	23
16		16	18	22	24
17		17-22	19-27	23	25
18				24-29	26-29
19			28-30	30-34	30

Table

Scaled Score Equivalents of Raw Score (Verbal) for 9 Years Children

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
1	0-2	0-13	0-5	0-4	0-4
2	3-9	13-14	6	5	5
3	10	15 to 18	7	6	6
4	11	19 to 24	8	7	7
5	12	24 to 26	9-11	8-10	8-9
6	13	27 to 29	12	11-12	10
7	14	30	13	13-14	11
8	15	31	14	15	12
9	16	32	15	16	13
10	17	33	16	17	14
11		34	17	18	15
12		35	18	19	16
13		36		20	17
14		37		21	18
15		38			19
16		39			20-21
17		40			22-23
18		41 to 51			24
19		51 above		23-34	25

Table

Scaled Score Equivalents of Raw Score (Non verbal) for 10 Years Children

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
1	0-10	0-3	0-10	0-4	0-13
2	11	3-8	11	5	14
3	12	9	12	6	15

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
4	13	10	13	7	16
5	14	11	14	8	17
6	15	14	15	9	18
7	16	15	16	10	19
8	17	16	17	11	20
9	18	17	18	12	21
10	19	18	19	13	22
11	20	19	20	14	23
12		20	21	15	24
13		21		16	25
14				17	26
15				18	27
16				19	
17				20	
18				21	28-29
19				22-30	30

Table

Scaled Score Equivalents of Raw Score (Verbal) for 10 Years old children

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
1	0-9	0-17	0-2	0-9	0-2
2	10	18 to 24	3	10	3-9
3	11	25 to 35	4	11	10
4	12	36 to 39	5	12	11
5	13	40	6	13	12
6	14	41	7	14	13
7	15	42	8	15	14
8	16	43 to 44	9	16	15
9	17	45	10	17	16
10	18	46	11	18	17
11		47	12	19	18
12		48	13	20	19
13		49	14	21	20
14		50	15	22	21
15		51	16	23	22
16		52	17	24	23
17		53	18	25	24
18		54	19	26	25
19		55 above	20-30	27-30	25-30

Table

Scaled Score Equivalents of Raw Score (Non verbal) for 11 Years Children

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
1	0-11	0-2	0-9	0-9	0-10
2	12	3-11	10	10	11
3	13	12	11	11	12
4	14	13	12	12	13
5	15	14	13	13	14
6	16	15	14	14	15
7	17	16	15	15	16
8	18	17	16	16	17
9	19	18	17	17	18
10	20	19	18	18	19
11	21	20	19	19	20
12	22	21	20	20	21
13	23	22	21	21	22
14	24	23	22	22	23
15	25	24	23	23	24
16	26	25	24	24	25
17		26	25	25	26
18		27	26	26	27
19	27-30	28	27	27	28-30

Table

Scaled Score Equivalents Raw Score (Verbal) for 11 Years Children

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
1	0-8	19 to 26	0-10	0-10	0-9
2	9	27	11	11	10
3	10	28	12	12	11
4	11	29 to 33	13	13	12
5	12	34 to 36	14	14	13
6	13	37 to 39	15	15	14
7	14	40	16	16	15
8	15	41	17	17	16
9	16	42	18	18	17
10	17	43	19	19	18
11	18	44	20	20	19

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
12	19	45	21	21	20
13	20	46	22	22	21
14	21	47	23	23	22
15	22	48 to 49	24	24	23
16	23	50 to 51	25	25	24
17	24	52	26	26	25
18	25	53	27	27	26-28
19	26-30	54 above	28	28	29

Table

Scaled Score Equivalents of Raw Score (Non verbal) for 12 Years Children

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
1	0-11	0-8	0-11	0-4	0-8
2	12-13	9	12	5-17	9-18
3		10	13	18	
4		11	14	19	19
5		12	15		20
6	14	13	16		
7		14	17		
8	15	15	18		
9		16	19		21
10	16	17	20		
11		18	21		
12		19			22
13		20			23
14		21			24
15	17				25
16				20	26
17				21	
18		22		22-25	27-33
19				26-34	34

Table

Scaled Score Equivalents of Raw Score (Verbal) for 12 Years Children

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
1	0-13	0-17	0-11	0-9	0-9
2	13-14	18 to 35	12	10	10
3		36	13	11	11
4	16	37	14	12	12
5		38	15	13	13
6		39	16	14	14
7		40 to 42	17	15	15
8		43	18	16	16
9		44	19	17	17
10	17	45	20	18	18
11		46	21	19	19
12		47	22	20	20
13		48	23	21	21
14		49	24	22	22
15		50	25	23	23
16		51	26	24	24
17		52	27	25	25
18	22-29	53 to 56	28	26	26
19	30	56 above	29	27	27

Table

Scaled Score Equivalents of Raw Score (Non verbal) for 13 Years Children

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
1	0-14	0-6	0-5	0-3	0-9
2	15-19	6-11	6-14	4-17	10-19
3	20	12	15	18	20
4	21	13	16		
5	22	14	17		21
6	23	15	18	19	
7		16			22
8		17			23
9	24	18		20	
10	25		19		25
11			20		26
12	26				27
13	27				

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
14	28	19		21	
15		20		22	
16					
17			21-22		28
18	28-32	21-25	23-25	23-28	29-30
19	33-36	25-30	25-30	28-30	31-34

Table

Scaled Score Equivalents of Raw Score (Verbal) for 13 Years Children

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
1	0-9	0 to 19	0-10	0-10	0-4
2	10-12	20 to 40	11-12	11-13	5-10
3	13	41		14	
4		42			11
5		43			12
6	14	44		15	13
7		45			14
8		46			15
9		47		16	16
10		48		17	17-18
11	15	49			19
12		50			20
13		51	13	18	
14	15	52	14	19-20	
15	16	53	15	21	22-23
16	17	54	16		
17		55	17		
18	18	56	18-21	21-23	24-28
19	19-30	56 above	22-30	24-34	29-34

Table

Scaled Score Equivalents of Raw Score (Non verbal) for 14 Years Children

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
1	0-13	0-5	0-5	0-9	0-15
2	14-20	6-12	6-15	10-17	16-20

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
3	21	13		18	21
4	22	14		19	22
5	23	15-16		20	23
6	24	17	16		24
7	25	18	17		25
8			18		26
9			19		27
10					28
11	26				29
12	27				30
13	28		20	21	
14		19	21	22	
15		20-21		23	
16		22			
17			23		
18	29-30	23-26	24-26	24-27	30-32
19	31-36	27-30	27-30	28-34	33-34

Table

Scaled Score Equivalents of Raw Score (Verbal) for 14 Years Children

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
1	0-11	0 to 19	0-10	0-7	0-3
2	12-13	20 to 40	11-12	8-14	4-11
3	14	41		15	12
4		42		16	13
5		43		17	14
6		44 to 45			15
7		46			16
8		47	13		17
9	15	48	14		18
10	16	49	16-17		19
11	17	50	18	18	20
12	18	51		19	21
13	19	52		20-21	
14	20	53 to 54		22	
15		55			22-24
16		56	19-20		25-27
17					
18		57 to 70	21-24	23-25	28-30
19	30	70 above	25-30	26-34	31-34

Table

Scaled Score Equivalents of Raw Score (Non verbal) for 15 Years Children

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
1	0-14	0-5	0-5	0-9	0-14
2	15-21	6-12	6-15	10-18	15-20
3	22	13	17	19	22
4	23	14	18	20	23
5	24	15	19	21	24
6	25	16	20		
7	26	17			
8					
9					
10		18			
11		19			
12		20			25
13	27	21	21-22		26
14	28	22	23		27
15	29	23	24	22	28
16				23	29
17				24-26	30
18	30-32	24-27	25-29	27-30	31-33
19	33-36	28-30	30	31-34	34

Table

Scaled Score Equivalents of Raw Score (Verbal) for 15 Years Children

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
1	0-3	0-19	0-2	0-9	0-4
2	4-13	20 to 40	3-12	10-14	6-12
3	14	41		15	13
4	15	42		16	14
5	16	43		17	15
6		44 to 45	13	18	16
7		46	14	19-20	17
8		47	15		18
9		48	16		19
10		49	17		20
11		50	18		21
12		51	19-20		22
13		52			23
14		53		21	24
15		54	21	22	25
16	17	55	22	23	26-27
17	18	56			
18	19-21	57 to70	23-25	24-26	
19	22-30	70 above	26-30	27-34	28-34

Table

Scaled Score Equivalents of Raw Score (Non verbal) for 16 Years Children

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
1	0-11	0-5	0-5	0-9	0-17
2	12-22	6-12	6-16	10	18-21
3	23	13-14	17-18	11	22
4	24	15-16	19	12	23
5	25	17	20	13	24
6	26	18		14	25
7	27			15	26
8				16-19	27
9				20	
10					
11					
12		19	21	21	28
13		20	22	22	29
14		21	23	23	30

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
15	28	22	24	24	31
16	29	23-24	25	25-26	
17	30-32		26	27	
18	-	26-28	27-29	28-30	32
19	33-36	29-30	30	31-34	33-34

Table

Scaled Score Equivalents of Raw Score (Verbal) for 16 Years Children

Scaled Score	Fluid Reasoning	Knowledge	Quantitative Reasoning	Visual Spatial Processing	Working Memory
1	0-9	0-19	1-9	0-10	0-8
2	9-13	20 to 45	10	11-14	9-14
3	14	46	11	15	15
4	15	47	12	16	16
5	16	48	13	16	17
6	17	49	14	17	18
7	18	50	15	18	19
8		51	16	19	20
9	19	52	17	20	21
10	20	53	18	21	22
11		54	19		23
12		55	20		24
13		56	21	22	25
14			22	23	26
15		57 to 60	23		27
16		60 to 65	24		28
17		66 to 69	25	24	
18	22-25	70	26-27	25-26	29-30
19	26-36	70 above	28-30	27-34	31-34

Table

Sum of Scaled Score (Non Verbal)

		Percentile	95% Confidence Interval (CI)		90% Confidence Interval (CI)	
NV Sum of Scaled Score	NVIQ	Percentile rank	Conf.	Int.	Conf.	Int.
5	68	5.1	67	70	67	69
6	70	6.1	69	70	69	70
7	71	7.1	70	71	70	71
8	72	8.1	71	72	71	72
9	73	9.1	72	74	72	73
10	74	10.2	73	74	73	74
11	75	11.2	74	75	74	75
12	75	12.2	75	76	75	76
13	76	13.2	76	77	76	76
14	77	14.2	76	77	76	77
15	78	15.2	77	78	77	78
16	78	16.2	78	79	78	79
17	79	17.2	79	79	79	79
18	79	18.3	79	80	79	80
19	80	19.3	80	80	80	80
20	81	20.3	80	81	80	81
21	81	21.3	81	81	81	81
22	82	22.3	81	82	81	82
23	82	23.4	82	83	82	83
24	83	24.4	82	83	82	83
25	83	25.4	83	83	83	83
26	84	26.4	83	84	83	84
27	84	27.4	84	84	84	84
28	85	28.4	84	85	84	85
29	85	29.4	85	85	85	85
30	86	30.5	85	86	85	86
31	86	31.5	86	86	86	86
32	86	32.5	86	87	86	87
33	87	33.5	87	87	87	87
34	87	34.5	87	88	87	87
35	88	35.5	87	88	88	88
36	88	36.5	88	88	88	88
37	88	37.6	88	89	88	89
38	89	38.6	89	89	89	89
39	89	39.6	89	89	89	89

40	90	40.6	89	90	89	90
41	90	41.6	90	90	90	90
42	90	42.6	90	90	90	90
43	90	43.6	90	91	90	90
44	91	44.7	90	91	90	91
45	91	45.7	91	91	91	91
46	91	46.7	91	91	91	91
47	92	47.8	91	92	91	92
48	92	48.7	92	92	92	92
49	92	49.8	92	92	92	92
50	92	50.8	92	93	92	93
51	93	51.8	93	93	93	93
52	93	52.8	93	93	93	93
53	93	53.8	93	94	93	94
54	94	54.8	94	94	94	94
55	94	55.8	94	94	94	94
56	94	56.9	94	94	94	94
57	95	57.9	94	95	94	95
58	95	58.9	95	95	95	95
59	95	59.9	95	95	95	95
60	96	60.9	95	96	95	96
61	96	61.9	96	96	96	96
62	96	62.9	96	96	96	96
63	97	64.0	96	97	96	97
64	97	65.0	97	97	97	97
65	97	66.0	97	97	97	97
66	97	67.0	97	97	97	97
67	98	68.0	97	98	97	98
68	98	69.1	98	98	98	98
69	98	70.1	98	98	98	98
70	99	71.1	98	99	98	99
71	99	72.1	99	99	99	99
72	99	73.1	99	100	99	100
73	100	74.1	100	100	100	100
74	100	75.1	100	100	100	100
75	101	76.1	100	101	101	101
76	101	77.2	101	101	101	101
77	102	78.2	101	102	101	102
78	102	79.2	102	102	102	102
79	102	80.2	102	103	102	103
80	103	81.2	103	103	103	103
81	103	82.2	103	103	103	103

82	104	83.2	103	104	103	104
83	104	84.3	104	104	104	104
84	105	85.3	104	105	104	105
85	105	86.3	105	106	105	106
86	106	87.3	106	106	106	106
87	107	88.3	106	107	106	107
88	107	89.4	107	108	107	108
89	108	90.4	108	108	108	108
90	108	91.4	108	109	108	109
91	109	92.4	109	110	109	110
92	110	93.4	110	111	110	111
93	111	94.4	111	112	111	112
94	112	95.4	112	113	112	113
95	114	96.5	113	115	113	114

Table

Sum of Scaled Score (Verbal)

V Sum of Scaled Score		Percentile	95% Confidence Interval		90% CI	
	VIQ	Percentile rank	Conf. Int.		Conf. Int.	
5	59	5.1	58	60	58	60
6	60	6.1	60	61	60	61
7	61	7.1	61	62	61	62
8	62	8.1	62	63	62	63
9	63	9.2	62	64	62	63
10	64	10.2	64	64	64	64
11	64	11.2	64	65	64	65
12	65	12.2	65	65	65	65
13	65	13.2	65	66	65	66
14	66	14.2	66	66	66	66
15	67	15.2	66	67	66	67
16	67	16.2	67	67	67	67
17	67	17.3	67	68	67	68
18	68	18.3	68	68	68	68
19	68	19.3	68	68	68	68
20	69	20.3	68	69	68	69
21	69	21.3	69	69	69	69
22	70	22.4	69	70	69	70

23	70	23.4	70	70	70	70
24	70	24.4	70	70	70	70
25	71	25.4	70	71	70	71
26	71	26.4	71	71	71	71
27	71	27.4	71	71	71	71
28	71	28.4	71	72	71	72
29	72	29.5	72	72	72	72
30	72	30.5	72	72	72	72
31	72	31.5	72	73	72	72
32	73	32.5	72	73	73	73
33	73	33.5	73	73	73	73
34	73	34.5	73	73	73	73
35	73	35.5	73	74	73	74
36	74	36.6	74	74	74	74
37	74	37.6	74	74	74	74
38	74	38.6	74	74	74	74
39	74	39.6	74	74	74	74
40	74	40.6	74	75	74	75
41	75	41.6	75	75	75	75
42	75	42.6	75	75	75	75
43	75	43.7	75	75	75	75
44	76	44.7	75	76	75	76
45	76	45.8	76	76	76	76
46	76	46.7	76	76	76	76
47	76	47.7	76	76	76	76
48	77	48.8	76	77	76	77
49	77	49.8	77	77	77	77
50	77	50.8	77	77	77	77
51	77	51.8	77	77	77	77
52	77	52.8	77	78	77	77
53	78	53.8	77	78	77	78
54	78	54.8	78	78	78	78
55	78	55.8	78	78	78	78
56	78	56.9	78	78	78	78
57	79	57.9	78	79	78	79
58	79	58.9	79	79	79	79
59	79	59.9	79	79	79	79
60	79	60.9	79	80	79	80
61	80	61.9	80	80	80	80
62	80	62.9	80	80	80	80
63	80	64.0	80	81	80	81
64	81	65.0	80	81	80	81

65	81	66.0	81	81	81	81
66	81	67.0	81	81	81	81
67	82	68.0	81	82	81	82
68	82	69.1	82	82	82	82
69	82	70.1	82	83	82	83
70	83	71.1	83	83	83	83
71	83	72.1	83	83	83	83
72	84	73.1	83	84	83	84
73	84	74.1	84	84	84	84
74	84	75.1	84	85	84	85
75	85	76.1	85	85	85	85
76	85	77.2	85	86	85	85
77	86	78.2	85	86	85	86
78	86	79.2	86	86	86	86
79	87	80.2	86	87	86	87
80	87	81.2	87	87	87	87
81	88	82.2	87	88	87	88
82	88	83.2	88	88	88	88
83	89	84.3	88	89	88	89
84	89	85.3	89	90	89	90
85	90	86.3	90	90	90	90
86	91	87.3	90	91	90	91
87	91	88.3	91	92	91	92
88	92	89.4	92	93	92	93
89	93	90.4	92	94	93	94
90	94	91.4	93	94	93	94
91	95	92.4	94	95	94	95
92	96	93.4	95	97	95	96
93	97	94.4	96	98	96	98
94	98	95.4	98	99	98	99
95	100	96.5	99	101	99	101

Table

FSIQ (Sum of Non verbal and Verbal Scaled Scores)

		Percentile	95% CI		90% CI	
Sum of NV and V Scaled Scores	**FSIQ**	**Percentile Rank**	**Confidence Interval**		**Confidence Interval**	
10	22	5.0	21	23	21	23
11	23	5.5	21	25	22	25
12	24	6.0	22	25	22	25
13	25	6.5	22	27	23	26
14	26	7.1	23	28	24	27
15	27	7.5	26	28	27	28
16	27	8.0	24	30	25	30
17	28	8.5	25	30	26	30
18	29	9.1	28	31	28	30
19	29	9.5	26	32	27	32
20	30	10.0	29	32	29	31
21	31	10.5	30	32	30	32
22	32	11.0	27	36	28	35
23	33	11.6	31	35	31	34
24	33	12.0	29	37	29	36
25	33	12.6	30	36	31	36
26	33	13.0	29	37	30	36
27	35	13.5	33	37	33	36
28	35	14.0	31	38	32	37
29	36	14.6	33	39	34	38
30	35	15.0	31	40	31	39
31	38	15.6	36	39	36	39
32	38	16.0	37	40	37	40
33	38	16.5	34	41	35	40
34	39	17.1	36	41	36	41
35	39	17.6	37	42	37	41
36	40	18.0	35	44	36	43
37	41	18.5	39	43	39	43
38	41	19.0	37	45	38	44
39	41	19.5	36	46	36	45
40	43	20.0	41	45	42	45
41	43	20.5	38	47	39	47
42	44	21.0	43	46	43	46
43	43	21.6	40	46	40	46

44	45	22.0	41	48	42	48
45	45	22.5	41	48	42	48
46	46	23.0	41	50	42	49
47	47	23.6	43	50	44	49
48	46	24.0	42	50	42	49
49	48	24.7	44	52	45	51
50	46	25.1	41	51	42	50
51	47	25.5	43	52	43	51
52	49	26.0	47	51	47	50
53	47	26.6	41	53	42	52
54	48	27.1	43	54	44	53
55	51	27.6	48	54	49	53
56	50	28.0	46	54	46	53
57	50	28.5	46	55	46	54
58	51	29.0	45	57	46	56
59	52	29.6	48	57	49	56
60	51	30.1	45	57	46	56
61	52	30.6	48	56	49	56
62	54	31.0	51	56	52	56
63	53	31.5	48	58	49	57
64	54	32.0	52	56	52	56
65	53	32.5	49	58	49	57
66	55	33.1	51	60	51	59
67	56	33.5	54	58	54	57
68	54	34.1	48	60	49	59
69	56	34.5	51	61	52	60
70	56	35.1	51	61	52	60
71	57	35.6	51	63	52	62
72	57	36.0	52	61	53	60
73	57	36.5	49	64	50	63
74	59	37.1	56	62	57	61
75	58	37.6	51	65	52	64
76	59	38.0	53	65	54	64
77	60	38.6	55	65	56	64
78	59	39.0	54	64	55	63
79	61	39.5	57	66	58	65
80	59	40.0	53	66	54	65
81	63	40.5	60	65	60	65
82	60	41.0	51	68	53	66
83	63	41.6	61	66	61	65
84	64	42.0	58	69	59	68
85	63	42.5	55	71	56	69

86	64	43.0	59	69	59	68
87	64	43.5	56	71	57	70
88	66	44.0	63	69	64	68
89	64	44.5	57	71	58	70
90	66	45.0	61	72	62	71
91	65	45.5	57	72	58	71
92	65	46.0	59	72	60	71
93	67	46.5	61	73	62	72
94	65	47.1	57	74	58	73
95	70	47.5	67	73	67	73
96	68	48.0	60	75	61	74
97	69	48.5	59	78	61	76
98	70	49.1	66	74	66	73
99	71	49.5	65	76	66	75
100	69	50.0	60	78	61	76
101	70	50.5	62	78	63	76
102	70	51.0	61	79	63	78
103	72	51.5	64	79	66	78
104	71	52.1	63	79	64	77
105	72	52.5	64	80	65	78
106	74	53.0	71	77	72	77
107	72	53.6	64	79	65	78
108	75	54.0	72	78	73	78
109	74	54.5	66	82	67	81
110	77	55.0	74	80	74	80
111	75	55.5	65	85	66	83
112	76	56.0	69	83	70	82
113	77	56.5	69	85	70	84
114	79	57.1	75	84	76	83
115	78	57.6	75	82	76	81
116	80	58.0	75	85	76	84
117	77	58.6	69	85	70	84
118	82	59.0	78	85	79	85
119	78	59.5	68	87	70	85
120	82	60.0	77	88	77	87
121	80	60.5	71	89	73	87
122	84	61.0	80	88	80	88
123	82	61.5	73	91	74	90
124	84	62.1	76	92	77	91
125	86	62.6	82	89	83	88
126	80	63.0	71	89	73	87
127	87	63.5	82	92	83	91

128	87	64.0	78	95	79	94
129	87	64.5	81	94	82	93
130	88	65.0	77	99	79	97
131	88	65.5	80	95	81	94
132	89	66.0	83	95	84	94
133	90	66.5	83	96	84	95
134	91	67.1	81	100	82	99
135	88	67.6	77	98	79	97
136	92	68.0	82	101	84	99
137	93	68.5	85	101	86	99
138	90	69.0	78	102	80	100
139	93	69.5	87	98	88	97
140	93	70.0	83	103	85	102
141	93	70.5	81	105	83	103
142	96	71.0	92	100	93	99
143	96	71.5	88	103	89	102
144	97	72.0	90	104	91	103
145	95	72.6	82	108	84	106
146	97	73.0	95	99	96	99
147	97	73.5	86	108	87	106
148	99	74.0	91	108	92	107
149	100	74.6	93	107	94	106
150	99	75.0	88	110	90	108
151	103	75.5	94	111	95	110
152	102	76.0	95	109	97	108
153	106	76.5	101	111	101	110
154	104	77.0	94	113	95	112
155	104	77.5	92	115	94	113
156	106	78.0	97	116	98	114
157	106	78.5	96	116	97	114
158	109	79.0	103	115	104	114
159	107	79.5	95	119	97	118
160	109	80.0	99	118	100	117
161	107	80.5	95	120	97	118
162	110	81.0	98	123	100	121
163	113	81.5	104	122	105	120
164	115	82.0	104	126	106	125
165	114	82.6	105	124	106	123
166	114	83.0	101	126	103	124
167	116	83.6	105	127	107	125
168	117	84.1	106	128	108	126
169	120	84.5	114	126	115	125

170	119	85.0	107	131	109	129
171	122	85.5	111	132	113	131
172	123	86.0	117	130	118	128
173	124	86.5	112	135	114	133
174	127	87.1	119	134	120	133
175	127	87.5	118	137	119	135
176	128	88.0	120	136	121	135
177	129	88.5	121	138	122	136
178	130	89.0	121	140	122	139
179	130	89.5	114	145	116	143
180	131	90.0	118	144	120	142
181	134	90.6	124	145	125	144
182	139	91.0	127	150	129	148
183	137	91.6	124	149	126	147
184	140	92.0	129	151	131	149
185	142	92.5	133	151	134	150
186	140	93.0	124	156	127	154
187	148	93.5	140	156	141	155
188	145	94.0	131	158	133	156
189	148	94.5	132	164	135	161
190	152	95.0	144	160	145	159
191	158	95.5	147	170	149	168
192	155	96.0	141	170	144	167
193	160	96.5	146	174	149	172
194	163	97.0	147	180	150	177
195	173	97.5	168	178	169	177
196	175	98.1	162	189	164	187
197	180	98.5	163	197	166	194

Table

IQ Range and Correlation in Non verbal Domain

Variable	Mean	Std. Dev.	Min	Max	Correlation	District
NVIQ_SB_BD_ALL	87.34	6.34	68	102	0.82	All
NVIQ_SB_BD_ALL	86.91	6.47	68	107		
NVIQ_SB_BD_M	86.27	6.78	68	99	0.79	Barisal
NVIQ_SB_BD_F	85.81	7.07	68	99		
NVIQ_SB_BD_M	89.56	5.34	68	102	0.81	Chittagong
NVIQ_SB_BD_F	89.53	5.24	73	107		
NVIQ_SB_BD_M	88.38	5.70	68	102	0.76	Dhaka
NVIQ_SB_BD_F	88.41	5.77	72	105		
NVIQ_SB_BD_M	86.81	6.49	68	102	0.84	Khulna
NVIQ_SB_BD_F	87.21	5.74	68	99		
NVIQ_SB_BD_M	87.29	6.04	68	98	0.80	Rajshahi
NVIQ_SB_BD_F	85.73	6.69	68	100		
NVIQ_SB_BD_M	85.73	6.81	68	99	0.87	Sylhet
NVIQ_SB_BD_F	84.79	6.89	68	98		

Table

IQ Range and Correlation in Verbal Domain

Variable	Mean	Std. Dev.	Min.	Max.	Correlation	District
VIQ_SB_BD_M_ALL	72.91	4.52	59	92	0.75	All
VIQ_SB_BD_F_ALL	72.54	4.64	59	89		
VIQ_SB_BD_M	73.21	4.68	59	91	0.75	Barisal
VIQ_SB_BD_F	72.43	4.96	59	86		
VIQ_SB_BD_M	73.10	4.59	59	89	0.76	Chittagong
VIQ_SB_BD_F	72.97	4.21	60	89		
VIQ_SB_BD_M	74.63	4.83	60	92	0.73	Dhaka
VIQ_SB_BD_F	74.26	5.06	59	88		
VIQ_SB_BD_M	72.93	4.38	59	89	0.70	Khulna
VIQ_SB_BD_F	73.33	4.10	59	88		
VIQ_SB_BD_M	72.09	4.18	59	85	0.79	Rajshahi
VIQ_SB_BD_F	71.19	4.23	59	81		
VIQ_SB_BD_M	71.50	3.79	61	89	0.74	Sylhet
VIQ_SB_BD_F	71.09	4.35	59	79		

Table

IQ Range and Correlation in FSIQ

Variable	Mean	Std. Dev.	Min.	Max.	Correlation	Districts
FSIQ_SB_BD_ALL	87.34	6.34	68	102	0.81	All
FSIQ_SB_BD_ALL	86.91	6.47	68	107		
FSIQ_SB_BD_M	109.60	9.78	88	139	0.94	Barisal
FSIQ_SB_BD_F	108.41	9.07	87	134		
FSIQ_SB_BD_M	112.68	8.87	93	148	0.95	Chittagong
FSIQ_SB_BD_F	112.25	8.36	93	152		
FSIQ_SB_BD_M	113.03	9.01	91	140	0.94	Dhaka
FSIQ_SB_BD_F	112.67	9.47	90	140		
FSIQ_SB_BD_M	109.79	8.94	90	148	0.92	Khulna
FSIQ_SB_BD_F	110.05	7.77	89	142		
FSIQ_SB_BD_M	109.27	8.00	90	128	0.89	Rajshahi
FSIQ_SB_BD_F	106.79	8.22	88	128		
FSIQ_SB_BD_M	107.24	8.22	88	145	0.92	Sylhet
FSIQ_SB_BD_F	106.09	8.20	88	128		

Appendix 12: Dutch-Bangla Bank Award

Dutch-Bangla Bank Foundation

Sena Kalyan Bhaban (5th Floor), 195, Motijheel C/A,
Dhaka-1000. Phone: 7176390-93 Ext: 172

DBBF/Fellowship/03/2006/836 SL # P-01

June 13, 2006

Nigar Sultana
D/O. Mafzal Ahmad
Lecturer, Dept. of Special Education,
Institute of Education and Research,
Dhaka University

Subject: Awarding of Dutch-Bangla Bank Foundation Fellowship.

Dear Nigar Sultana,

We feel honored to award you a fellowship under the fellowship program of Dutch-Bangla Bank Foundation (reference to Letter # DBBF/Fellowship/02/2006/ dated May 28, 2006). The amount of the fellowship will be Tk.5,000/-(Taka five thousand) only per month for 01 (one) year initially (from July 2006 to June 2007).

The fellowship money will be credited to your bank account on **quarterly basis**. You are, therefore, advised to report to the **Dhanmondi Branch** of **Dutch-Bangla Bank Limited**, **House # 580-A/1 (1st Floor), Road # 8, Dhanmondi R/A, Dhaka**, and open a Savings Bank Account in your name **within June 20, 2006 to June 27, 2006. You should furnish the account number to the office of DBBF** on or before **July 05, 2006**. It may be mentioned here that the fellowship money for the month of **July 2006** amounting to **Tk.5,000** (Taka five thousand) only will be credited to your bank account on or after **August 16, 2006**. Then **Tk. 10,000/-**(taka ten thousand) only for the next **02** (two) months will be credited on or after **October 07, 2006**. Henceforth the rest fellowship money will be credited to your bank account on **quarterly basis**.

The fellowship may be renewed from year to year basis at the sole discretion of the Foundation, for the total duration of your course **(not exceeding 03 years)** subject to satisfactory research performance duly certified and recommended by the Research supervisor/Departmental Head. **You should apply for renewal (in plain paper) within 30 (thirty) days after completion of the present period of fellowship along with the supporting documents mentioned in our previous letter. Failure of which will lead to auto cancellation of the fellowship for that particular period.**

The terms and conditions of the Fellowship will remain same as communicated through our previous letter # DBBF/Fellowship/02/2006 dated May 28, 2006.

Thanking you,

Yours truly,

Sd/-
Medical Consultant

C.C. **Head of Department / Dean.** He is requested to kindly inform this office if there is any adverse report against the student or he/she discontinues studies during the tenure of the Fellowship

Dr. Mozammel Hossain Khan
Medical Consultant

NB. 1) A photocopy of this letter should be submitted to the office of Dutch-Bangla Bank Foundation along with your Account Number for disbursement of fellowship money in your account.

Alfred Binet (1857 – 1911)

Stanford BINET

Intelligence Scales

Gale H. Roid (1943)

SB5 Test Kit