DEVELOPMENT, VALIDATION AND STANDARDIZATION OF A GROUP VERBAL INTELLIGENCE TEST IN URDU FOR ADOLESCENTS

DEVELOPMENT, VALIDATION AND STANDARDIZATION OF A GROUP VERBAL INTELLIGENCE TEST IN URDU FOR ADOLESCENTS

By

SYED SAJJAD HUSSAIN

Dr. Muhammad Ajmal
National Institute of Psychology
Centre of Excellence
Quaid-i-Azam University
Islamabad-Pakistan
2001

DEVELOPMENT, VALIDATION AND STANDARDIZATION OF A GROUP VERBAL INTELLIGENCE TEST IN URDU FOR ADOLESCENTS

By

SYED SAJJAD HUSSAIN

A thesis submitted to the

NATIONAL INSTITUTE OF PSYCHOLOGY
Centre of excellence
Quaid-I-Azam University
Islamabad-Pakistan

in the partial fulfillment of the requirements
For the degree of

DOCTOR OF PHILOSOPHY
In
PSYCHOLOGY

2001

DEVELOPMENT, VALIDATION AND STANDARDIZATION OF A GROUP VERBAL INTELLIGENCE TEST IN URDU FOR ADOLESCENTS

By

SYED SAJJAD HUSSAIN

Approved by

Supervisor

External Examiner

Director
National Institute of Psychology

CERTIFICATE

Certified that Ph.D. Dissertation " **Development,Validation and Standardization of a GroupVerbal Intelligence Test in Urdu for adolescents"** prepared by Syed Sajjad Hussain has been approved for the submission to Quaid-I-Azam University, Islamabad.

(SUPERVISOR)

CONTENTS

LIST OF TABLES

Before administering the sub-tests for psychometric characteristics in main study, it was decided to give the title of ***Sajjad Verbal Intelligence Test in Urdu (SVITU)*** *to the newly developed test. Average time for SVITU and sub-tests was calculated. SVITU was administered to a sample of 535 candidates of PMA Long Courses who had completed grades 12 and were in the age range of 17 to 20 years. The sample was taken from the Army Selection and Recruitment Centres of Pakistan. The purpose of the main study was to find out different psychometric characteristics of the test.*

The reliability of the SVITU was determined by Kuder Richardson method, Split-half method, and test-retest method. The results suggest high reliability of the test both in terms of internal consistency and temporal stability of the results.

The validity of the test was determined by four procedures. The construct validity of the test was studied by correlating it with Army Intelligence Test. The concurrent validity was determined by correlating it with the college marks of the students in their past annual examination. Other two techniques were internal consistency of the sub-tests and grade differentiation. All the four indices established the evidence of high validity of the test.

Percentile norms, Z-scores and T-scores were computed for grade 12 boys only.

Before administering the sub-tests for main study, average time for IIT and sub-tests was calculated. The final test was administered to a sample of 535 candidates of PMA Long Courses who had completed grades 12 and were in the age range of 17 to 20 years. The sample was taken from the Army Selection and Recruitment Centres of Pakistan. The purpose of the main study was to find out different psychometric characteristics of the test.

The reliability of the IIT was determined by Kuder Richardson method, Split-half method, and test-retest method. The results suggest high reliability of the test both in terms of internal consistency and temporal stability of the results.

The validity of the test was determined by four procedures. The construct validity of the test was studied by correlating it with Army Intelligence Test. The concurrent validity was determined by correlating it with the college marks of the students in their past annual examination. Other two techniques were internal consistency of the sub-tests and grade differentiation. All the four indices established the evidence of high validity of the test.

Percentile norms and T-scores were computed for grade 12 boys only.

INTRODUCTION

CHAPTER-1

INTRODUCTION

Historical Background

Intelligence is not only a topic of immense importance that is becoming central to psychology as a discipline; it is also one of the extensively explored areas in the realm of psychometrics. The primary scientific basis for asserting that general intelligence exists and that individual differences in intelligence have some impact on behavior is the observation that there is a positive correlation between every reliable measure of mental ability. Intelligence is a construct conceptualized differently by various psychologists. Boring (1923) defined intelligence as what the intelligence tests measure. Sternberg (1986) carried this viewpoint and operationally defined intelligence in terms of the way it is measured. Operational definition of intelligence has two shortcomings (Sternberg, 1986). First, it is circular i.e, intelligence tests are devised to measure intelligence, not to define it. Second, operational definition blocks further progress in understanding the nature of intelligence. In contrast, a real definition is one that seeks to tell the true nature of the concept being defined (Sternberg, 1986). Most of the definitions of intelligence appeared in an early symposium, *Intelligence and its measurement* (Thorndike, 1921).

Spearman (1904, 1923) defined intelligence as the general ability, which involves the process to educe the relations and correlates. Binet and Simon (1905) viewed intelligence as the ability to judge well, to understand well, and to reason well. Terman (1916) defined intelligence as the capacity to form concepts and to grasp their significance. Brown (1921) defined intelligence as the ability to perform certain types of tasks. Pintner (1921) explained intelligence as the ability to adapt adequately to relatively new situations in life. Intelligence, according to Thorndike (1921), is the power of good responses from the point of view of truth and fact. Thurstone (1921) defined intelligence as the capacity to inhibit instinctive adjustments, flexibly imagine different responses, and realize modified instinctive adjustment into overt behavior. According to Wechsler (1939), intelligence is the aggregate or global capacity of the individual to act purposefully, to think rationally, and to deal effectively with the environment. Humphreys (1971) viewed intelligence as the entire repertoire of the acquired skills, knowledge, learning sets, and generalization tendencies considered intellectual in nature and that are available at any one period of time. Piaget (1972) explained intelligence as a generic term to indicate the superior forms of organizations or equilibrium of cognitive structuring used for adaptation to the physical and social environment. Eysenck (1979) considered intelligence as the error free transmission of information through the cortex. Sternberg (1985a, 1986) defined intelligence as the mental capacities to automatize information

processing and to emit contextually appropriate behavior in response to novelty. Intelligence is usually associated with the abilities to learn quickly, to adapt to new situation, to use abstract reasoning, to understand both verbal and mathematical concepts and to perform the tasks in which relationship is grasped (Bruno, 1986). Intelligence, according to Gardner (1993), is the ability or skill to solve problems or to fashion products which are valued within one or more cultural settings.

Evolution of Intelligence Testing

History of intelligence measurement began with the publication of Sir Francis Galton's early work, *Hereditary Genius* (1869), and flourished with the formation of the Psychometric Society, particularly with the issue of first journal *Psychometrika* in 1935, and is still under the process of development and refinement. Till 1935, major problems in the study of mental abilities were identified, the basic methodologies were developed, and tests of mental ability came into wide use.

Intelligence testing developed within a societal and educational sphere like Darwin's formulation of the theory of evolution and *Social Darwinism* of Hofstadter (1944). Focus of attention of *Social Darwinism* was hereditary differences in the adaptability of individuals to the new demands of an industrialized and technological society. Galton was of the view that acuity of senses and the reaction time could be used as important indices of mental

ability. Galton in his *Hereditary Genius* (1869) worked on the problem of delineating the characteristics of genius persons and suggested that genius tends to run in families. From 1884 to 1890, Galton in his anthropometric laboratory developed not only mental tests but also introduced the notion of correlation. However, it was the American psychologist James McKeen Cattell (1890), who introduced the term mental test. Working with Wundt at Leipzig, Cattell became interested in individual differences in reaction time, sensory discrimination, word association, and other simple mental tasks (Cattell, 1886). He used tests of all these capacities and advocated that these tests might be used as appropriate predictors of scholastic achievement of the college students (Cattell and Farred 1896).

Binet (1890) was the first to trace the development of certain abilities over the age range of childhood and early adolescence. A criterion for the acceptance of a task into his series was that it shows not only progressive increase in performance over these ages but also shows consistent relationships with degrees of mental deficiency over these ages. The credit of measuring intelligence in an appropriate way goes to the Binet-Simon scale (1908), which became the basis for the development of individual intelligence tests in many countries. It was recognized by several American psychologists (Otis,1918) that some of the mental tasks of Binet-Simon Scale could be adapted for group testing to meet a practical need to classify the recruits during World War-1. In this way, the first tests of group intelligence called the Army Alpha Examination and Army Beta Examination were developed.

Scores on both of these tests were scaled in terms of mental age, the concept given by Binet.

By 1935, group mental testing had become widely adapted basing on the view that intelligence was a more or less a unitary trait. Most tests of intelligence produced a single score for an individual, expressed as an *intelligence quotient* or some derivative of the intelligence quotient. The common view was that scores on intelligence test revealed innate capacity of the individual to think in abstraction and to adapt to the requirements of technological society. Intelligence tests were valued as scientific measurements to the extent that they showed high degrees of reliability and validity. Commercial publishers were interested for developing multifactor batteries for wide use in secondary schools. One of the most popular and extensively used of such batteries is Differential Aptitude Tests (DAT), published by the Psychological Corporation (Bennett, Seashore, & Wesman, 1947). This battery offers separate scores on verbal reasoning, numerical ability, abstract reasoning, space relations, mechanical reasoning, clerical speed and accuracy, and language usage.

Brigham (1968) was the first to develop systematic item analysis procedures using large samples. He was also the first to stress the importance of test security in the administration of admission tests. Brigham avoided calling his test an *intelligence test* and emphasized the difference between verbal and mathematical aptitudes.

Psychometrics and Theories of Intellectual Abilities

Mental tests consist of a sample of almost infinite domain of the tasks that an individual might be required to perform in real life, however, these samples of tasks appearing in mental tests are constrained in at least two important ways:

1. Only those kinds of tasks can be included in a psychological test that are small enough to be completed successfully by at least some examinees in a relatively short period of time, and that they must not involve any complicated apparatus for stimulus presentation or for the examinee's response.

2. The tasks in a mental test are selected in such a way that successful performance of a task requires one or more kinds of intellectual competence or potential.

Given the characteristics of mental test, an adequate theory is required, that is, a theory of what such tests measure and how well they measure. An adequate theory gives a satisfactory account of at least the following:

a) The extent to which the tasks can yield reliable measures of the characteristic behaviors of individuals, to the limit that the

measures derived from the tasks are in fact reliable.

b) The extent to which successful performance on each task, or group of tasks, reflects some particular kind of cognitive competence.

c) The source or cause of the individual differences observed with respect to any scale of identified cognitive competence.

The Role of Test Theory in Ability Testing

Test theory was initiated in the early period of the mental test movement by Spearman, Brown, Thorndike, Kelley and others and was concerned chiefly with the reliability of measurement. Concern for exactly how this quantification took place was expressed only in the development of various item-analysis techniques that would ensure that all items or tasks involved in obtaining a given measurement would tend to be consistent with one another and thus contribute effectively to the measurement. Reliability was defined as the extent to which observed measurements reflected a true underlying score. This approach is known as classical test theory.

Kuder and Richardson's (1937) formulation for estimating reliability from the *internal consistency* of the items and Cronbach's (1951) derivation of

coefficient alpha led to determine the reliability of the test. These reliabilities were more accurate and more covenient because they could be based on the administration of a single test. Test theorists began to look at the effects of different distributions of item difficulties on the measurement characteristics of total test scores. Attention was devoted to the attenuation paradox, whereby it seemed that maximum reliability, but reduced validity for measurement of an underlying ability, was attained when the distribution of item difficulties was *peaked,* that is, when all item difficulties were near 50 percent difficulty.

A related issue that has been studied extensively is that of how item scores can be weighed to produce optimal measurement characteristics of total scores (Stanley &Wang, 1970). Test theorists also showed concern in the development of *tailored testing,* that is, a procedure for optimizing the efficiency of a mental test by the sequential items of a wide range of difficulty.

Stanley (1971) believes that ANOVA or *components of variance* approach provides much greater precision and flexibility in the estimation of reliability. A still further development of this approach, called, *generalizability theory* is mainly due to Cronbach and his associates (Cronbach, Gleser,Nanda, & Rajaratnam,1972; Cronbach, Rajaratnam, & Gleser,1963); it provides procedures for estimating the reliability of a test that would be derived from a *universe* of measurements.

A persistent theme in the development of mental tests is a concern with the way in which successful versus unsuccessful performance on a given item or task relates to an underlying *latent trait* continuum of ability. Binet (Binet & Simon, 1905) assumed that a one to one correspondence could be established between graded tasks and mental ages, but experience with Binet-type scales revealed that the correspondence was far from exact. What is needed is a model for the probability of an examinee to pass an item or task successfully as a function of the characteristics of the items or task, principally, its *difficulty* and *discriminating power* and the examinee's position on an underlying continuum of ability. Such a function has been called the item-characteristic function (Lord & Novick, 1968).

Guttman (1965), and Carroll (1945) described a model, according to which if an examinee passes an item of a given level of difficulty, he or she passes all items of that difficulty and if an examinee fails an item of a given level of difficulty, he or she also fails all items of that difficulty. Loevinger (1947) suggested a measure of the extent to which a test approximates perfect homogeneity. Wherry and Gaylord (1943) were the first to point out that the interpretation of reliability coefficients determined by Kuder-Richardson, or similar methods, required the assumption that all the items measure a single ability or composite of abilities. A perfectly homogeneous test is one in which Cronbach's coefficient alpha is equal to the theoretical reliability of the test.

Views about Intelligence

Before developing a test of intelligence, one should study the efforts of Galton, Cattell, Binet, and many others. Galton was interested in the characteristic of genius; Binet was concerned with distinguishing grades of mental deficiency and contrasting them with the abilities of the average child. Binet was successful in devising a scale that distinguished grades of mental deficiency.

Brown and Thomson (1921) proposed, the *sampling theory* of intelligence. According to this view, any mental task samples a wide variety of mental operations; the correlation between two tests is therefore a function of the amount of overlap between the sets of operations.

Thorndike (1927) presented the general view of nature of intelligence. Thorndike saw intelligence as a general capacity for forming bonds or connections among ideas, concepts, and so on. Persons of high intelligence are those who have the capacity to form a large number of bonds and have had the opportunity to do so. Thorndike viewed intelligence as having both hereditary and environmental components.

Spearman (1950) developed the *two-factor* theory of intelligence whereby each test is regarded as measuring one *general* factor in common with all the other tests and, in addition, a *specific* factor that is unique to that test. His major writings that bear on this subject are *The Nature of Intelligence*, The *Principles of Cognition* (1923), *The Abilities of Man (1927)*, and *Human Ability (1950)*. Spearman conceded that all tests could not be interpreted in terms of a two-factor theory. His later writings recognized the appearance of *group* factors and even two other specialized *general factors*, preservation and oscillation alongside the original g-factor. Spearman held that the *general* factor is central and supreme in all tests of intelligence.

Guilford (1967; 1988) does not accept the existence of a general intellectual factor. Rather, he proposes that each combination of a specific operation, a specific type of content, and a specific type of product defines a unique type of intelligence. According to Guilford's structure of intellect model, there are 180 different types of intelligence, defined by all possible combinations of six operations, five types of contents, and six types of products (Murphy & David Shofer,1998.p 25).

Sternberg (1977, 1980, 1981b, 1982, 1985) has proposed a different structure of the domains of the intelligence. His theory notes that mental activities can be broken down into components, which are used to acquire the information, to carry out specific mental tasks, and to evaluate general mental

processes. The theory suggests that adaptation to one's environment is a critical measure of intelligence (Sternberg, 1984).

Early theories of the cognitive processes involved in intelligence tended to focus on elementary information-processing operations, or the ways in which we perceive, encode, store, and retrieve information Jensen (1980). More recent theories have worked from a broader base. For example, Sternberg's (1985) triarchic theory of intelligence features three subtheories. His componential theory, which is concerned with *how* intelligent behavior is generated, deals with the basic cognitive processes involved in knowledge acquisition and performance. His contextual theory deals with *what* behaviors are intelligent in what contexts. His experiential theory deals with *when* a given behavior is intelligent.

Carroll (1993) presented three-stratum model of cognitive ability according to which, at the most general level, there is a *g* factor, responsible for stable differences in the performances on the wide variety of cognitively demanding tasks. At the next level (the broad spectrum), there are a number of areas of ability, which imply that the rank ordering of individual's task performance will not be exactly the same across all cognitive tasks, but rather will show some clustering. The broad abilities in Carroll's model include (1) fluid intelligence, (2) crystallized intelligence, (3) general memory ability, (4)

broad visual perception, (5) broad auditory perception, (6) broad retrieval ability, and broad cognitive speediness. Some people will do well on the broad range of memory tasks and others will do well on the a broad range of tasks of cognitive speediness. These broad ability areas can be characterized in terms of number of more specific abilities (the narrow spectrum). The narrow spectrum includes (1) induction, (2) language development, (3) memory span, (4) spatial relations, (5) sound discrimination, (6) word fluency, and (7) perceptual speed. According to Carroll, tests designed to measure *g* and the tests designed to measure more specific aspects of intelligence both have independent place. It is up to the researcher to choose the ability area in the hierarchy according to the purpose of testing rather than by personal preferences.

The Debate Over *g*

Factor '*g*' is central to the theories of Spearman, Vernon, Carroll, and the like. On the one hand, there is sufficient evidence that there are broad and general differences in the performances of the people on cognitively demanding tasks and that a general factor emerges from analyses of virtually any set of cognitive tests. People's standing on the general factor is probably

the best single predictor of their performances in school, on the job, a other settings where cognitively demanding tasks must be performed. Once is taken into account, considering additional more specific abilities, does not seem to help much in predicting performances in these settings (Ree & Carreta, 1994;Ree & Earles, 1991a, 1991b, 1992;Ree, Earles, & Teachout, 1994;Murphy, 1996). From a pragmatic point of view, it is hard to doubt the utility of the *g* factor.

On the other hand critics of *g-ocentric* theories of intelligence (e.g., McClelland, 1993; Sterngerg,1984,1985;Sternberg& Wagner,1993)observed that while concentrating much on the *g* factor, one is likely to lose sight of the key questions such as why a person is intelligent, or what intelligent people do, or how they do it. Murphy (1996) suggests that models of intelligence that focus on *g* are most useful for pragmatic purposes like prediction of future performances, whereas models that focus on facets of intelligence other than *g* may be more useful for structuring research on the nature and origins of intelligence.

Adolescent's Intelligence

Testing has been oriented toward the school child and college student. At these levels, the test constructor can draw on the large common pool of experiences that have been organized into academic curricula. Most

intelligence tests measure how well the individual has acquired the intellectual skills taught in our schools and predicts how well he is prepared for the next level in the educational hierarchy. Tests for adults, including the Wechsler scales, draw largely on this identifiable common fund of experience.

As the individual grows older and his own formal educational experiences recede further into the past, this fund of common experience may become increasingly less appropriate to assess his intellectual functioning. Adult occupations are more diversified than childhood schooling. The cumulative experiences of adulthood may stimulate a differential development of abilities in different persons.

Whether intelligence test scores rise or decline with increasing age in adulthood, depends largely on what experiences the individual undergoes during those years. Longitudinal studies of adults show larger age increments in score among those individuals who have continued the education longer (Campbell, 1965; Harnqvist, 1968; Husen, 1951; Lorge, 1945; Owens, 1953). Similarly, persons whose occupations are more *academic* in content including verbal and numerical abilities, are likely to maintain their performance level or show improvement in intelligence test scores over the years, while those engaged in occupations emphasizing mechanical activities or interpersonal relations may show a loss.

Intelligence and Social Context

It is not fair to assume that human intelligence can exist as a trait rather than a product of social and cultural context. An individual who is used to specific social context will appear to be deficient when tested in an unfamiliar context. Reasoning is not only embedded in the mental representations of physical contexts but even more so in the social context. According to Vygotsky (1962), intelligence is an identifiable human characteristic that is an organizing principle of social thought and activity. There is no such thing as intelligence as an individual characteristic separate from social context. Sternberg (1985) has put forward a triarchic theory of intelligence which asserts that intelligence is a purposeful and goal oriented behavior consisting of two general skills; the ability to deal with novel tasks and to learn from experience. Intelligence depends on acquiring information processing skills and strategies to solve problems. Intelligence cannot be understood outside a socio-cultural context. What may be relevant in one culture may not be in the other culture. For an example, one of the items comprising Comprehension subscale of the Wechsler's Adult Intelligence Scale is: ***If you were lost in the forest in the day time how would you go about finding your way out*** is a very relevant question in the USA where tens of thousands of people get lost every year. However, there are many countries in the world where it would be absurd to test intelligence with such a question.

Testing and describing mental abilities of people require far more sensitivity in the social context in which abilities are developed and expressed. Test situations are rarely simple tasks. They involve other people. Although Raven's Progressive Matrices have been described as assessing intelligence in the contextualised way and is said to be culture-free (Keating, 1980), even then a culture-free test cannot be described as a pure cognitive task isolated from social context.

Academic and Non-academic Intelligence

The academic view of intelligence focuses mainly on the individual's cognitive processes and also on a strong underlying belief that those who meet the challenges of the academic world can also meet the challenges of the world. But the real world challenges may require certain strategies, styles, modes of cognizing and skills that may not be related to the academic performance. Hence, it is possible for an illiterate to be intelligent who can show in the real world all the hall marks of intelligent behavior, like motor mechanic or plumber etc.

A unitary concept of intelligence assumes that intelligence is an innate ability that can be summoned into different cognitive abilities. It is the innate potentiality for achievement in areas like abstract reasoning, verbal analysis, creative expression, quantification, visio-spatial organizations, and so on.

Environmental opportunities enable the individuals to develop these potentialities. Therefore, being proficient in one area means being equally proficient in other areas but most of us are not equally gifted in all these potentialities nor do we have equal opportunities to develop these abilities evenly, even if we have the same opportunities. The fact is that most of us can perform on complex levels only in some of the areas for which we have the potentialities and we have had the opportunities to emphasize these potentialities and we have the motivation to take advantage of such opportunities. Thus we do not function in an equally complex manner in all aspects of life or in all the tasks that are set to assess intelligence. The concept of generality of intelligence across tasks and across situations is not substantiated, if more of a theoretical assumption rather than a practical reality.

Sternberg (1983) believes in the existence of multiple intelligence, An implication of this view of intelligence is that the individual possessing a lower IQ may exhibit a high level of cognitive abilities in the non-academic setting, whereas those with a higher IQ will display less cognitive abilities in non-academic matters because of remaining environmental challenges leading to differential employment of the underlying intelligence. Hence, IQ scores do not predict social behavior, but only a level of ability.

Intelligence and educational achievement

There exists a vast correlational literature on intelligence tests as predictors of later measures of educational achievement (Cattell & Butcher, 1968; Lavin, 1965; Jensen, 1980.). The accumulated evidence supports five related points that need to be accounted for in further research:

1. The average correlation between general mental tests and educational achievement measures is typically found to be about, .50. Intelligence and achievement measures are closely related but not equivalent psychologically. As individuals grow and education progress, the relation is reduced.

2. General ability and achievement are thought to differentiate with age through childhood and adolescence. The best evidence supports the view that ability and achievement differentiate as a result of education (Anastasi, 1970).

3. At least in the public school years, correlation between intelligence measures taken at one point in time and achievement measures taken at a later point in time tend to be higher than when the measures are taken in the reversed time - order (Crano, 1974; Crano, Denny, & Campbell, 1972). The

theory of fluid and crystallized intelligence (Cattell, 1963, 1971; Horn 1976, 1982) suggests that the individual in learning experiences, including formal education, to produce verbal-crystallized ability, invests the presumably more native, fluid analytic ability. The growth curve for fluid ability appears to precede that for crystallized ability across the childhood and young year.

4. The more general tests, of both fluid and crystallized abilities are highly correlated and more involved in the organization of human abilities than are the more specialized, peripheral skills and abilities (Guttman, 1965;1969 Jensen, 1970; Snow, 1980b). It is the more general, central ability measures that correlate most highly with educational achievement measures, especially when the latter reflect more generalized achievement criteria rather than performance in specific or special subject matters. Although fluid and crystallized abilities are at times difficult to distinguish in this regard, verbal crystallized ability relates more to achievement in relatively conventional, familiar instructional setting (e.g, lecture, recitation reading, discussion, etc), whereas fluid -analytic ability relates to achievement in which novel problem solving or adaptation to unfamiliar instruction methods and material is involved (Snow, 1981).

5. There appears to be marked variation in ability-achievement relation around the central tendencies usually obtained. This depends not only upon the ability and achievement measures used, the school level (s) and subject (s) in which achievement is assessed, and the educational environments sampled. Such variation suggests that the relation of intellectual measures to educational achievement is moderated by environmental variables. Potentially, greater understanding ability-achievement relations could be gained by analyzing, perhaps even manipulating experimentally, the instructional environmental variables that influence such relations.

Verbal Abilities

Tests of verbal abilities (reasoning, comprehension, and expression) have been the most widely used of all intelligence measures. They are the best single index of an individual's overall intelligence in predicting future performance. In the advanced countries of the world, comprehensive testing program cannot be envisaged without including verbal tests.

For groups who have not attained a certain level of language proficiency, tests of ability will yield a greatly distorted picture of learning

potential. Considerable effort has been devoted to the construction of tests to measure the general mental capacity of an individual without the use of any language content whatever. As a replacement for words, abstract shapes, drawings, mazes, and numerous other essentially spatial configurations have been applied, and a wide variety of entirely nonverbal tests, including all of the so-called culture-fair measures, has been developed.

Inductive Reasoning Ability

Sternberg (1977, 1979; Sternberg & Gardner, 1983; Sternberg & Nigro, 1980) has proposed a theory of inductive reasoning that he has been applied to analogies, series, completions, classifications and metaphorical understanding. The theory of information processing specifies seven components that seem to be involved in inductive reasoning. The components are (1) encoding, by which the individual recognizes the terms of the problems and accesses attributes of analogy terms that are stored in semantic memory and that might be relevant for task solutions; (2) inference, by which the individual figures out the relation between the first two terms of the analogy (e.g. that a lawyer provides professional consulting services to a client); (3) mapping, by which the individual figures out the higher order relation between professional services (4) application, by which the individual takes the relation inferred between the terms in the first half of the analogy as mapped to the third term in the second half of the analogy and uses this

relation to generate an *ideal* completion; (5) comparison, by the which the individual compares each answer option in multiple- choice analogies, to the ideal, and decides which is better in the sense of more closely resembling the ideal; (6) justification in which the individual decides whether the preferred answer option is close enough to be entertained; and (7) response, by which the individual communicates his or her choice of an answer (for example, the individual might circle an answer or press a button indicating his or her choice of the preferred answer).

The theory can also be applied to other kinds of induction problems, for example, in the series involving digits, the individual must encode the terms of the problems, infer the relation between each successive pair of given digits, apply this relation to generate the next digit in the series, compare each of the two answer options to the generated option, possibly justify the chosen option if it does not correspond exactly to one of the given options. Mapping is not required in the series completion because the problem did not require any recognition of higher order relations. The series completion problems thus require only the component processes required by the analogy.

In a set of experiments, Sternberg's theory was tested developmentally (Sternberg & Nigro, 1980, Sternberg & Rifkin, 1979). In the Sternberg - Nigro study, verbal items were used and the investigators looked at use of word association as well as of reasoning in the solution of the items. For

example, in the analogy, TREE: ANIMATE: PENCIL: (a) PAPER, (b) INANIMATE. Paper has the greater associative relation to PENCIL, even though the correct answer to the analogy is INANIMATE. Children ranging in educational levels from grade two to college receive either one of two kinds of schematic-picture analogies (Sternberg- Rifkin) or verbal analogies (Sternberg- Nigro). In these experiments, the most interesting data proved to be qualitative data illustrating the functioning of components.

To sum up, Sternberg's theory of inductive reasoning contains two parts: a theory of information processing and a theory of response choice. The theory of information processing specifies the processing components used by individual as they actually solve an induction problem. It includes processing components that is inferring relations, applying relations, and mapping relations. The theory of response choice, based on a theory proposed by Rumelhart and Norman (1981) seeks to predict individual's response choices in inductive reasoning. It uses a representation of a multidimensional psychological space to make its predictions.

Inductive reasoning problems are characterized by an absence of a single, logically certain response. One might see in an inductive reasoning test the series completion: 1,2,3, 4 ---. One's task is to complete the series with the number 5, and indeed, if this problem appeared in an intelligence test for children, it is virtually certain that 5 would be viewed as the correct answer.

One of the first theorists of general intelligence, Spearman (1923), used analogies as the prototypes for intelligent performance. Spearman exemplified ability to perceive second-order relations, or relations between relations that has served as the touchstone marking in the transition between concrete and formal operations in Piaget's (1972) theory of intelligence. Analogies, since they require the ability to perceive relations between relations for their solution, can serve as a useful measure for distinguishing concrete-operational from formal operational children (Sternberg & Rifkim, 1979).

Indigenization

Indigenous approach to the social sciences recognizes the fact that psychological phenomena are both meaning and context dependent. Indigenous psychology emphasizes the need to provide a rich description of a psychological phenomenon as the first step in the scientific endeavor. According to Webster's dictionary, indigenous is defined as native, not introduced directly or indirectly into a particular land even through historical record or scientific analysis. An indigenous approach is native originating or developing naturally in a particular land or region. Indigenous psychology can thus be defined as a psychological knowledge that is native, that is not transported from any region, and that is designed for its people. In other words, indigenous psychology is understanding, rooted in a particular socio-cultural context, it emphasizing the use of natural taxonomies. Each culture

needs to develop its own indigenous understanding of its own culture.

The majority of the researchers in our country have followed a culturally inappropriate social science which distorted interpretation of behavior, diverted our attention from the relevant and key social variables and resulted in research that matched little with our social problems and national priorities, To address these issues we need the indigenization of discipline of psychology, i.e., a research that reflects back on the culture in which behavior is studied rather than a discipline that is imported from and primarily addresses developed world models (Adair, Puhan & Vohra, 1993).

In the area of adaptation and development of psychological assessment tools, test adaptation and development in the future may be changed on two lines. First, we have to practice psychometry or psychodiagnostic within the framework of our cultural norms, social values, folklore and traditions. Second, we should educate our professionals on modern lines and acquire the necessary expertise to see new methods, models and advance statistical measures to adapt, to develop and to practice psycho diagnostics.

The standardized psychological tests borrowed from the West can be adapted for local use to a greater or lesser extent. The technical know how, the experience and the expertise gathered through the adaptation and use of Western tests should culminate in the development and increasing use of

indigenous tests that may be more sensitive to the local demands.

The degree of adaptation of a foreign test can range from use *without change* to a test that has been substantially altered i.e., translated and items reworded and or deleted to accommodate cultural difference. Intermediate stages of increasing adaptation include minor word changes, translation only, and translation plus minor word changes. In adaptation of tests, it is simply stated by the majority of our psychometricians that a test was *adapted*, without specifying the magnitude and kind of changes.

As far as test construction is concerned, the first step is to write items for the test. However, before writing items test specifications are to be prepared to avoid imbalances and disproportion of item coverage. An exhaustive trait sampling is required of what has to be measured.

The second important step in test construction involves selection of the sample. The sample should be big and representative of the population in order to generalize the results and to perform multivariate and other complex analyses to stabilize the results. In our multi-ethnic society care should be taken to equally represent different ethnic groups in the national sample. Each ethnic sample included in the national sample should be big enough to undergo different statistical analyses to determine the similarity and dissimilarity in the distributions of test scores and test profiles.

One problem relates to the form of a language as well as the categories of relationships in a specific culture which do not exist in other cultures, hence the language lacks such concepts or words. For example, the word *cousin* in English language does not indicate all types of different cousins for each of which we have a different word in Urdu. The word *cousin* cannot be translated with all its functional significance. There are other words and concepts that do not have any equivalence in another language. For Example, *I love Sind* when translated in Urdu literally will have a different shade of meaning. Love in our cultural context has certain meaning but in translation it will be *Mujhay pasand hay,* that is *I like.*

Even bilingual scholars face such difficulties. For example, an Urdu sentence, *Mein mussibat mein ghir gaya thaw* was translated by an English man who knew Urdu well, as *I was surrounded by misfortunes.* It shows how delicate the differences become when sentences are translated in a foreign language (Sechrest et. al. 1972).

Another difficulty relates to the translation of responses of the subjects. After all, the local people who are not going to give you the responses in English would give the responses in the local language. For example, if one goes to a village of Punjab and prepares an instrument with adequate translation, the responses would be either in Urdu or in Punjabi. How is he

going to translate those responses into English in order to make them comparable.

Another technique used for translation of a foreign made test is called back translation, which is being used in modern social science research. Back translation means that the items are translated into Urdu and again in English and then give this English translation to some one, who has not seen the Urdu version, to translate it into Urdu and then compare the original Urdu with this translated Urdu and vice versa. The items often are quite different and even funny when translated into English because of these difficulties.

The problem of attempting equivalence in research instruments is not entirely a linguistic one. There are certain other conceptual equivalences, which are still not possible to achieve. There may be difficulties if one does not find the equivalent concept in other language. It is not only the linguistic equivalence but also the conceptual equivalence that is required. For example, that a profession may be male in one culture and female in the other, like Vendors in Philipine are mostly females while in Pakistan probably they are both. To translate the sentence, *I want to be a vendor*, the translation may confuse the sex role associated with the job. Therefore, unless one knows the culture well, the equivalence may not be easy to achieve.

Another problem faced, sometimes, pertains to certain concepts that

have different connotations in different socio-cultural contexts, for example, *teacher*. A teacher in England is of a very high social prestige not comparable to that in many other countries including Pakistan. Can we translate a teacher as *Ustad*, a university professor as *Muddaris*, or a madrasa teacher as *Maulvi*? We have so many types of teachers who have different social backgrounds.

Group Tests of Intelligence

Developed initially during WW1 to meet the practical demand for classification of US Army recruits, the group tests became increasingly popular, with the passage of time. They transcended their original role, and own definitions of intelligence, mental ability, or cognitive functioning.

The concept of intelligence underlying all the group tests that enjoyed popularity in the 1920s was of a *general* ability manifested in an individual's performance across a wide variety of cognitive tasks. This concept is nearly the same as that of Spearman's '*g*'. This view of intelligence influenced the selection of content, the scoring and interpretation of results of the early group tests. Although the authors of tests included a variety of tasks in their tests, such as vocabulary, analogies, arithmetical reasoning, etc, these varieties of content were chosen because of their supposed suffusion with g, having little regard for their mutual independence or possible diagnostic utility. The use of varieties of tasks was intended to facilitate *sinking different shafts* to permit a

comprehensive sampling of the hypothesized general ability. Scores on the several parts were added to yield an overall measure of mental ability, expressed as a mental age and an IQ. Even those tests that offered verbal and nonverbal, or language and non-language features treated these as different modes of combining the two measures into a single IQ.

Group tests make greater demands on understanding written and spoken language than do individual tests, and thus may place at a disadvantage to subjects either of foreign language background or of deficient reading skills. It is also true that group tests are always time limited and may thus have an element of speediness, however, almost all the group tests have been designed to minimize the effect of speed on performance. Obviously the group test does not generate a protocol as the individual test is likely to, though with respect to the great majority of subjects and of uses made of test results, this limitation is not serious. The reliance on multiple-choice items has risen to the criticism that guesswork or chance unduly influences results of group tests. For the further criticism that they place too great a premium on the recognition process rather than the supposedly more creative processes tapped by the individual tests; empirical evidence on the reliability and predictive validity of the two types of tests suggests that these differences are not of major significance.

The role of psychologists in the area of psychometry is being realized in Pakistan. Test development programs are being demanded and initiated. Increasing attention is being given to tests of achievement, personality, aptitude, and intelligence. Practical problems and issues in evolving a systematic approach for establishing test development programs for intelligence in Pakistan have been identified and strategies for dealing with such issues and problems have been suggested.

Ahmed (1986) developed Academic Self-Concept Scale (ASCS) as a measure of academic assessment of self-concept for use in high schools. The scale showed an Alpha coefficient of .89. ASCS was found to have validity of .37 with high school grade and predictive validity of .39 with marks for matriculation examination 10 months later, and significantly differentiated between students who were academically poor and those were average.

Ahmed (1987) adapted and translated in Urdu, the California Psychological Inventory (CPI). After administration of Urdu and English versions on bilingual subjects, the results showed sufficient similarity between the two versions. The overall psychometric evaluation showed sufficient credence to the further use of CPI in Pakistan and to research on predictive and classificatory issues.

Ain (1985) studied the validity of Cattell's Culture Fair Test of Intelligence (CFIT) for Pakistani children. It was found that while CFIT scores are unrelated with age, they are strongly related with grade, achievement test and teacher ratings. Retest reliability was also found to be significantly high.

Ansari (1976) developed an abbreviated version of Wallach-Kogan Creativity Test, that was correlated with AH5 Verbal Part and Standard Progressive Matrices. It was found that while the intelligence is equally related to the achievement test of Lower and Higher Cognitive objectives, the creativity measures show significantly higher correlations with achievement test of Higher Cognitive Objectives.

Hasan (1981) assessed the effects of bilingualism on the performance of Pakistani school girls on tests of verbal intelligence and reasoning. The results indicated that bilingualism was significantly related to poor performance in verbal intelligence and reasoning.

Ismail and Mahmood (1986) administered Raven's Standard Progressive Matrices to 300 students to study the effect of sex and social class. A significant difference was found between the performances of three different social classes. No significant difference was found between males and females.

Israr (1985) tried out thirteen Piagetian tasks on a sample of 360 primary school children from all over Pakistan. The findings show that grade1 children are at early concrete-operational stage and grade 3 children are at mid concrete operational stage while grade 5 children are at late concrete operational stage.

Israr (1982) studied the patterns of language used in parent child interaction in Pakistani society. Any analysis in terms of its relationship with cognitive development is not attempted.

Kausar and Sheikh (1986) studied the effect of immediate and delayed knowledge of results on subsequent performance in verbal learning. The results indicate that immediate knowledge of results tended to favor the subsequent performance of all subjects, with no significant difference between the performance of males and females.

Khan (1987) studied the relationship between ideational fluency and intelligence among academically gifted and average students.. It was found that intelligence tests as compared to Creative Thinking Test did not discriminate well between the academically gifted and average group of students.

Najam, Andrabbi, Malik, and Ghaznavi (1990) investigated the

differences in some personality variables and intelligence, which could be attributed to the difference in birth orders. As regards intelligence, the youngest scored highest on the test of mental ability.

Riaz (1979) studied the relationship between intelligence and creativity. The resuls demonstrated a substantial degree of indepenenceof these two variables. However as hypothesized, a positive and significant relationship was found between intelligence and academic achievement as well as the creativity and academic achievement. The author argues that creativity is a mode of cognitive behaviour independent of intelligence, but each contributes positively to academic achievement. Riaz (1975) established the reliability and validity of Raven's Standard Progressive Matrices for Pakistani population.

Moghni and Riaz (1984) investigated the relationship between motives, study habits, and academic performance. The findings demonstrated a high correlation between study habits and academic performance, whereas the relationship between *n* achievement and academic performance was low.

Ansari (1988) developed a questionnaire for assessing the study problems of Pakistani high school students in Peshawar. The questionnaire was found to have moderately high level of alpha reliability between .74 and .84 for various groups.

Ansari and Iftikhar (1988) conducted a study of the validity of Raven's Standard Progressive Matrices for urban and rural school children. The results showed that RSPM is useful as test of intellectual performance for the urban school children but their utility for the rural children is not so certain.

Ansari, Tariq and Iftikhar (1990) developed and validated Educational Ability Test Level 5. The test purports to evaluate the current status of a student in terms of a broad range of cognitive educational objectives including his/her ability to recall, comprehend, reason and analyze material that the student comes across his/her environment in the school and outside the school. The internal consistency and test-retest reliability were also found to be satisfactory: ranging from .87 to .92 for various groups.

Imam and Munaf (1988) administered Raven's Standard Progressive Matrices on 66 grade 5 female students. A significant difference in intellectual performance was found among the first, second and third born children.

Hussain (1992) developed a Group Verbal Intelligence Test in Urdu for High School students with an aim to facilitate the students of those high schools where the medium of instruction is Urdu.The test comprised of two subtests: Vocabulary Test and Numerical Reasoning Test. Validity indices for Vocabulary Test and Numerical Reasoning Test were .82 and .58 respectively.

The reliability indices for Vocabulary Test and Numerical Reasoning Test were .86 and .82 respectively. The reliability index of the total test was .88. Percentile norms were computed for grade 10 of RawalPindi City.

Syed (1993) developed a nonverbal test of intelligence for Pakistani urban primary school children. This test comprises of two subtests: Block Design Test and Picture Completion Test. Validity indices of the two sub-tests were .896 and .893 respectively and were significant at .01 level.

Naheed (1993) developed a verbal intelligence test in Urdu for Pakistani Urban Primary School Children in the age range of 5 to 11 years. This test consisted of two subtests: Vocabulary and Arithmetic. The validity indices for Vocabulary, Arithmetic and total test were .79, .56, and .85 respectively. The reliability indices for Vocabulary, Arithmetic and total test were .84, 82 and .86 respectively.

Gardezi (1994) developed a non verbal intelligence test for adolescents of grade 10 in the age range of 15 to 17. The test comprised of four subtests viz Series, Analogies, Classification and Matrices.The reliability index was .82 for KR-20, and .77 for split half categories respectively. The construct validity of the test was computed by correlating it with a Group Verbal Intelligence Test in Urdu (GVITU) developed by Hussain (1992), which was .76. Percentile norms were developed separately for boys and girls of grade10.

Israr (1988) studied psychological interpretation of mathematical learning problems among the urban students of grades 6^{th} to 8^{th} of secondary schools. The results show that there are more problems in learning mathematics at grades 6 and 7 as compared to the grade 8. Mathematical ability and general ability are difficult to distinguish from each other. The findings also revealed that mathematical questions presented in narrative forms are more difficult to learn as compared to the non-worded questions.

Khalid (1986) investigated differential schooling effects on cognitive development and intellectual ability of primary school children selected from government and private schools. The results showed that there was no significant correlation between school type and cognitive development indicating that the differences between cognitive levels of the children of different schools are largely attributed to the socio-economic status of the parents and probably with little contribution from the school.

Khalique (1982) studied audience effect on the speed of vocabulary writing. The results indicate that the subject's speed of writing English words correctly lowered down before the audience as compared to their speed in solitude. Moreover, the speed of writing enhanced when others were engaged in the same task as compared to working before audience.

In an early study, Khalique (1981) explored the effects of physical presence of others on the speed of verbal learning of introverts and extroverts. No significant difference is observed in lonely situation, whereas significant differences were reported in learning while others were present.

Khalique (1983) studied the effect of language similarity between the experimenter and the subject on the speed of verbal learning. Results show that speed of learning was greater when instructions were given in Punjabi as compared to the situation when the instructions were given in English. The primary recency effect of the language did not appear to have any bearing on performance.

Mahmud (1990) developed and validated Educational Ability Test for Pakistani pre-school children. The test consisted of 56 items and covering six areas: Visual Matching, Reasoning, School Language, Quantitative Concepts, AuditoryMemory, and Rhyming. The Retest reliability indices and KR 20 computed for the total test were reported to be .82 and .90 respectively. These reliability indices demonstrate the stability and homogeneity of the test. Significant differences were found in the urban-rural samples, while gender differences were indicated only in rural sample.

Zoofashan(1982) developed a computer model for statistical analysis of progressive matrices. The objective of the project was to develop a

computer model for a psychologist with no or a little knowledge of computer to check the reliability and validity of progressive matrices in a sample of Pakistani children of both sexes, belonging to urban and rural areas.

Limitations in cross-cultural testing

Vernon (1969, 1979) discussed numerous problems involved in testing intelligence cross-culturally. These problems affect the test scores such that they may not accurately reflect the intelligence of the subjects. These problems include unfamiliarity with the test situation, lack of motivation, anxiety, excitement, and suspicion of the tester when the psychologist is of different race. In some cultures, there may be difficulties with particular types of items or materials. Lack of test sophistication is another factor that adversely affects the test scores. Moreover, there may be problems in understanding the test instructions.

There may be other handicaps in some cultures like lack of varied perceptual and kinesthetic experience, restricted linguistic stimulation, lack of interest in formal education. Limited environmental stimulation, little schooling and emphasis on rote learning in school, are the most important environmental factors which Vernon (1979) lists as likely to affect intelligence test scores. A problem raised by Vernon (1979) concerns the handicaps caused by poor medical care and nutrition, which can lower intelligence test scores.

While these are restricted to cultures other than the West, they are more prevalent in society.

If a test is developed for a particular culture so that the items are meaningful, the task familiar, and the instructions comprehensible, and the examinees understand the nature of the test and its purpose, it is then possible to test intelligence with some degree of accuracy. Many cross-cultural psychologists (e. g., Berry and Dasen, 1974) consider that this is indeed the case and that it is not a meaningful question to compare the intelligence of members of different cultures because different qualities are valued differently in different cultures which makes assessment of crystallized intelligence difficult.

Vernon (1979) suggests that comparison between the cultures might be possible if there is equal access to education, equal freedom from physical disabilities, equal familiarity with the test and freedom from test anxiety and equal valuation of the skills involved in the test in the two groups. It might be said that this rules out most cross-cultural comparisons, especially those between advanced and third world countries. So comparison on tests of intelligence between different cultures are dubious even though accurate measurement within cultures is possible if tests are properly developed.

METHOD

CHAPTER- 2

METHOD

Aims and Objectives of the Study

This study was aimed at achieving the following goals:

a) Development, validation and Standardization of a Group Verbal Intelligence Test in Urdu for Adolescents.

b) To determine the reliability of the test.

c) To establish the validity of the test.

d) To develop norms of the test for Pakistani adolescents.

Development, Validation and Standardization of a Group Verbal Intelligence Test in Urdu for Adolescents.

The focus of intelligence testing has always been directed chiefly towards the school child and college student. Nearly each intelligence test measures the intellectual ability acquired through educational curricula and is used to predict how well the student is prepared for the next level in the educational hierarchy. The main function of intelligence tests is to measure the individual differences. Assessment of intellectual levels and classification of the children with reference to their abilities, the identification of the intellectually retarded on the one hand and the gifted on the other, diagnosis of

academic failures, and the selection of the applicants for different jobs are the main uses of the intelligence tests. Intelligence testing constitutes an important part of the total personnel selection program in Pakistan. A closely related application of intelligence testing is found in the selection and recruitment of military personnel in Armed Forces of Pakistan.

In every learning situation, language is the essential medium by which new information is acquired, and a verbal test assesses this process more closely than will ever be possible with a nonverbal procedure. Verbal items are bounded more firmly in skills that the examinees have practiced before, than are the essentially artificial tasks that comprise most of the nonverbal approaches. A test that is fair to all cultures is not likely to yield a meaningful ability measure.

Another reason for choosing verbal tests is that the substitution of pictures or symbols for words increases the problems of test adaptation. Probably as a function of the education methods, words are a more effective medium for testing than are symbols or pictures.

It was decided to develop only verbal intelligence because at the same time when this research was under progress, another researcher was developing nonverbal intelligence test nearly on the same target population. Hence it was assumed that both independent researches will supplement each other in future for all purposes of selection and recruitment.

Verbal ability test can be adapted for use with marginally qualified group. With foreign social scientists working in a different culture, the question of language becomes quite significantly crucial. If it is the English, it would be foreign to local subjects and if it is the local language, the foreign expert will not understand it.

Cattell (1943, 1973) had called attention to a possible distinction between fluid intelligence and crystallized intelligence, the former representing basic capacity and the later representing abilities acquired through learning, practice, and exposure to education. Both fluid and crystallized intelligence could be regarded as having associated with them a larger number of factors representing narrow abilities at the level of Thurstone's Primary Mental Abilities.

The two development streams can be labeled *crystallization* and *fluidation,* which correspond to the two kinds of ability, respectively (Snow, 1981). Crystallized intelligence represents the organization of more formal educational experience into functional cognitive systems applicable in educational situations. Fluid ability, on the other hand, is thought of as analytic reasoning, particularly where flexible adaptation to novel situations is required and where crystallized ability offers no particular advantage.

Crystallized intelligence represents previously constructed assemblies of performance processes retrieved as a system experience in the past, whereas

fluid intelligence represents new assemblies, or the flexible reassembly of performance processes needed for more extreme adaptations to novel situation. Both functions develop through exercise, and perhaps both can be understood as variations on a general production system. It is possible that the crystallized assemblies result from the accumulation of many *fast-process* intentional learning experiences. Whereas the facility for fluid assembly and reassembly results more from the accumulation of slow-*process*, incidental learning experiences. Both kinds of intelligence are relevant to education. Fluid ability pertains more to learning performance with new or unusual instructional methods of content. Crystallized ability is more relevant in the progression of familiar situations.

In later versions of the fluid and crystallized theory of intelligence, Cattell (1971) and Horn (1994) have proposed a hierarchical, interlocking model of intelligence with fluid and crystallized components at the top of their model.

Vernon (1960, 1965), in hierarchical model of intelligence, suggested a general factor '*g*' that pervades all tests of intelligence and that can be broken down into two broad categories, called *major group factors*. These two major factors are verbal-educational and spatial-motor, which can be further divided into minor group factors, verbal and spatial factors.

Factor analysis of abilities carried out by Hakstian and Cattell (1974) yielded two factors of intelligence, which are, fluid ability (g_f) and crystallized ability (g_c). Fluid ability closely resembles Spearman's g in that it involves the capacity to perceive relations and correlates with the materials that minimize individual differences in education and acculturation. Crystallized ability on the other hand, involves the same capacities but as they are realized within any given culture.

Typical intelligence tests designed for use with adolescents or adults, measure largely verbal abilities and to a lesser extent, they also deal with numerical and other abstract symbols (Anastasi, 1990). As intelligence is not a single unitary ability, the term is commonly used to cover the combination of abilities required for survival and for advancement within a particular culture (Anastasi, 1986b)

Since majority of the adults in Pakistan have a better command of Urdu language as compared to English which is the medium of foreign tests available in Pakistan, the researcher felt a strong need to develop a standardized indigenous intelligence test in Urdu language that can be a better predictor of our population especially those studying in Urdu medium institutions. The proposed test was designed to assess vocabulary, verbal reasoning, numerical reasoning, and information through four subtests.

Vocabulary Test

The ability to read rapidly with good comprehension is a crucial aspect of *intelligent* behavior in any literate society. Vocabulary test presents written stimulus words and asks the respondents for their meanings. Vocabulary in all forms has long been recognized as the most accurate, stable and general measure of mental abilities or intelligence. Within the WAIS-R, Vocabulary has the highest correlation with Verbal IQ across all age groups included in the standardization sample. Many brief forms of the WAIS-R make use of Vocabulary as one of the best single measure of intelligence. House and Lewin believe that WAIS-R Vocabulary subtest is an excellent measure of the verbal ability (Newmark, 1985).

The ability to define words is not only one of the best measures of intelligence, but it is also the most stable and least deteriorating aspect of intelligence (Rapaport et al., 1968). Vocabulary tests are included in nearly all individual tests devised to measure verbal intelligence. As vocabulary scores provide an estimate of general verbal intelligence that is relatively stable and independent of deterioration, it can be used to evaluate the base line intelligence, that is, what a person's intellectual capacity probably was prior to an emotional illness, brain injury or trauma.

Another reason for selecting vocabulary test is that it is one of the most preferred estimate of intelligence when only a brief test is desired (Jenson, 1980). According to Cattell (1971), vocabulary test loads highly on fluid ability at an early age but gradually shifts over to crystallized ability. Vocabulary scores are not just dependent on educational experience, they do demand reasoning and mental processing of past.

Verbal Reasoning Test

Verbal Reasoning Test is a measure of ability to understand concepts framed in the words. It is aimed at the evaluation of the subject's ability to abstract or generalize and to think constructively. The analogy items are appropriate for the measurement of reasoning ability. The particular type of analogy item recommended and constructed for the test is an efficient double-ended analogy in which both the first term of the first pair and the last term of the second pair are missing. The examinee is asked to choose the one pair that best completes the analogy from among four pairs of words given as the options. This style of items is especially useful to provide a measure of reasoning. Another advantage of the analogies is that the content of the items may be varied in as many ways as desired. The words used in these items may come from history, geography, literature, or any other content area. The items assess the subject's knowledge and ability to abstract and generalize relationships inherent in that knowledge.

Numerical Reasoning Test

Numerical reasoning is concerned with the ability to invent solutions to problems. Although simply algebraic problems involve numbers, the main ability being measured is not that of numerical computation. However, the necessity to work the problems in your head, the common *math anxiety,* in much of the general population, and the timed nature of the test combine to make this test mildly stressful for many people. In order to solve the problem, the subject must concentrate, analyzed the nature of the problem, grasp the principle by which each problem can be solved, remember the partial results until the problem is solved and finally the answer reported. The general reasoning factor also appears in the items concerning serial completion, in which the subject is required to supply the response in the patterned series of letter or digits. There is an element of discovery in all tests that measure the factor of general ability, hence Numerical Reasoning test is also loaded on g-factor, measuring fluid and crystallized intelligence.

Information Test

The information subtest involves both intellective and non-intellective components including the ability to comprehend instructions, follow directions, and provide a response. The items of this subtest consist of factual inquiry questions sampling the subject's general fund of knowledge. The item

content is selected to represent the sort of background information that a developmental exposure to the Pakistani culture will bring within the common knowledge of most citizens. Information is believed to reflect the acquired knowledge that correlates with formal education. However, a pattern of generalized reading across diverse topics will also yield a high information score regardless of formal education.

Information along with Vocabulary, is usually seen as one of the best measure of general ability among WAIS-R subtests and has the second highest correlation with the verbal IQ and the full scale IQ's. It consistently loads highly on the first verbal factor identified in most common factor and principal component factor analyses. In developing tests for use within a single familiar country, psychometricians have studied the differences in the knowledge of the intra-country subcultures, such as the black and the white, the rich and the poor, or the urban and the rural sections of the population.

Still, there are two major points to be made about the use of test content that is consistent with the knowledge available in a particular cultural setting. The first is that there are few instances in which the nature of the knowledge elements that should be included is immediately *obvious* and can be determined without special study. A more systematic procedure for generating suitable test content usually must be developed. The second related point is that the knowledge elements of a standard test that in fact are

inappropriate in our cultural setting may not be at all *obvious*. Some elements, though strange to the examinees, may not affect their performance because they are not central to the problem being solved and changing such elements simply because they are strange may be a needless extravagance in situations in which economy in test development is important.

Steps of Test Construction.

For the construction of a Group Verbal Intelligence Test in Urdu for adolescents, standard procedure was adopted. Following were the steps followed for the development, validation, and standardization of the proposed intelligence test.

Item Generation

Items of the test, being the scoring unit, require lot of efforts and precision for its construction. It was decided to include multiple-choice form of items as it is free from many of the weaknesses inherent in other forms of items. It is adaptable to a wide variety of contents and is being used to measure complex abilities and fundamental understandings about basic knowledge. Since there is only one correct response to a multiple-choice item, the difficulty and subjectivity of scoring that plague the short-answer form are avoided.

Vocabulary Test

The source of items of Vocabulary Test was Urdu textbooks used at intermediate level of education in Pakistan. While constructing the test items, one word from every fifth line of every fifth page of Urdu textbooks was selected. The initially generated item pool consisted of 250 items for Vocabulary Test, which included only common nouns (Isams) and adjectives.

Verbal Reasoning Test

The Verbal Reasoning Test is a measure of ability to understand concepts framed in words. This test is designed to evaluate an individual's ability to think, extract, and generalize rather than focusing on comprehension of vocabulary. The item format used in this test is an efficient double-ended analogy in which both the first term of first pair and last term of second pair are missing. The examinee is asked to choose from among the four pair of words, the pair that best completes the analogy. The correct pairs consisted of the first part for the missing part of the first pair and second part for the missing part of the second pair. This type of format is a useful measure of logical reasoning as it is slightly complex without being tricky. The words used in these items require general awareness about various professions, religion, geography, science and related context areas. This test aims at assessment of an individual's ability to abstract and generalize relationships

Vocabulary Test

The source of items of Vocabulary Test was Urdu textbooks used at intermediate level of education. While constructing the test items, one word from every fifth line of every fifth page of Urdu textbooks was selected. The initially generated item pool consisted of 250 items for Vocabulary Test, which included only common nouns (Isams) and adjectives.

Verbal Reasoning Test

The Verbal Reasoning Test of IIT is a measure of ability to understand concepts framed in words. This test is designed to evaluate an individual's ability to think, extract, and generalize rather than focusing on comprehension of vocabulary. The item format used in this test is an efficient double-ended analogy in which both the first term of first pair and last term of second pair are missing. The examinee is asked to choose from among the four pair of words, the pair that best completes the analogy. The correct pairs consisted of the first part for the missing part of the first pair and second part for the missing part of the second pair. This type of format is a useful measure of logical reasoning as it is slightly complex without being tricky. The words used in these items require general awareness about various professions, religion, geography, science and related context areas. This test aims at assessment of an individual's ability to abstract and generalize relationships

inherent in their knowledge. A total of 64 items were generated for Verbal Reasoning Test.

Numerical Reasoning Test

To generate the items of Numerical Reasoning Test, the basic concepts and principles of arithmetic, such as addition, subtraction, multiplication, division, percentage etc, were used, which are present in the textbooks recommended for intermediate level. In the construction of items, main emphasis was laid upon perceptual clarity, comprehension and application rather than mere reproduction of factual knowledge. In other words, the items included in the Numerical Reasoning Test evaluate ability of individuals to reason with numbers and deal with quantitative material in an intelligent way. The items generated for Numerical Reasoning Test were 70 in number.

Information Test

Items of Information Test were selected from the books of general information covering different concepts, purely from the Pakistani culture. Efforts were made to select the items of information, which seem to be the part of common pool of experience of the population so that the generated items should load on the g factor. The items generated for this subtest were 52 in number.

Selection of Distracters

For finding the distracters of the items, all the items of Vocabulary Test, Verbal Reasoning Test, Numerical Reasoning Test, and Information Test were administered to 50 students of class 12 of PAF Intermediate College, Rawalpindi and 50 candidates of 105 PMA Long Course. All the items were of free-response type. This procedure was used to obtain good distracters from the wrong answers actually given by the students. Three plausible responses were selected as the distracters for each item of the four sub-tests. Correct answers to items of Vocabulary Test were taken from Urdu Dictionary (Feroz ul Lughat). Correct answers to the items of Verbal Reasoning Test, Numerical Reasoning Test, and Information Test were obtained from the teaching staff of PAF College, Rawalpindi. The researcher also consulted some popular books on general information for instance, Current Affairs, Who is Who? and What is What ?. An effort was made to change the serial position of the correct answer to each item randomly with a purpose to eliminate the element of guessing by the subjects.

An effort was made to construct the items in such a manner that people who know the answer to the item always chooses the correct answer and the people who do not know the answer, chose randomly among the possible distracters. This means that some people may guess correctly despite the fact that every distracter is equally popular.

After finalizing the format of the items of the four sub-tests, instructions were written for the sub-tests separately and also for the total test. All the instructions were simple and brief. Examples were also used to clarify the instructions.

Language of the test

To make the test partially indigenous, Urdu being the national language was used. The purpose of using Urdu language was to generate the items maximum on the basis of the concepts and terms used in our own culture and which are in the common pool of experience of the target population.

Initial Editing and arrangement of the Items

Once the items were ready in their initial form, they were presented to four psychologists for their expert opinion and advice. They were requested to grade each item on the basis of its perceptual complexity and face validity, on five-point scale. They were requested to tick either of 1,2,3,4 or 5 for every item. Grade 1 being the easiest and 5 being the most difficult. In the light of their views and grading, items of Vocabulary Test were reduced from 250 to 95, the items of Verbal Reasoning Test were reduced from 64 to 40, items of

Pre-Tryout

Before the items were tried out, they were assembled into tryout forms. By a Pre-tryout, it is meant that the preliminary administration of the tentative tryout to small sample is carried out with the purpose to discover gross deficiencies in the individual items.

The item pool was administered to a group of sixty students, ten each from classes 11 and 12 of three different colleges. The students were introduced to the purpose of the study and then were given instructions to take the test. The results of this administration revealed that no modification was required and it was decided to use the same format for the actual administration.

First Experimental Tryout

The first experimental tryout of the proposed sub-tests (first draft) was carried out to serve the following purposes.

1. Item analysis of the items of each sub-test to determine the following item characteristics:-

a) Internal consistency.

b) Difficulty level.

c) Discriminatory power.

2. Revision of sub-tests in the light of the above mentioned information.

Sample

The sample consisted of 200 candidates of 104 PMA Long Course appearing at Army Selection and Recruitment Center, Rawalpindi. This sample was selected simply because these candidates constituted a representative group of different areas of Pakistan fulfilling the predetermined criteria of having intermediate qualification and falling within 17-20 years of age. Random sampling procedure was used for the selection of candidates to have a balanced and more equally distributed group with reference to their various socioeconomic backgrounds.

Procedure

Procedure adopted for the First Try-out was as follows:

Test Administration

First draft of the proposed four sub-tests was administered to eight groups of candidates, each comprising twenty-five candidates making the total of 200. Efforts were made to reduce test anxiety by presenting the test as a research instrument rather than an achievement test. The candidates were seated in a relaxed atmosphere for the reduction of test anxiety. They were assured of the confidentiality of the results. At the beginning of each

administration, the instructions printed on the cover page of each booklet, were clearly read out to the subjects. The candidates were instructed to attempt all the questions and not to leave any question, as there was no time limit for taking the test. They were encouraged to ask questions, if any, regarding the instructions, examples and recording of their responses. Since no group raised any question, it was assumed that instructions were clear and unambiguous and were properly understood by the candidates.

Scoring and Analysis of Data

The computer using SPSS package carried out all scoring and analyses of the data for determining the psychometric properties of the items.

Item Analysis

Item analysis was the major objective of this experimental tryout. Each item was analyzed to find out its effectiveness, considering the role it could play in future revision, and for developing the final version of the test.

For item analyses of the items of the proposed sub-tests, three indices were calculated:

a) Internal Consistency of the items: To see the relevance of each item with the test.

b) Index of item difficulty: To see the percentages of subjects who gave correct answer to the item.

c) Index of item discrimination: To see if the test discriminate between high and low achievers on the sub-test.

To determine the internal consistency of the items and examine their relevance with the test, item-total correlation was calculated, because item total correlations are directly related to the reliability of the test (Nunnaly,1982). To determine the difficulty level and discriminatory power of the items, method of contrasted groups was used. As suggested by Cureton (1957), 27 percent subjects were selected from top and bottom groups on the basis of their test scores to make two extreme groups.(see Murphy & David shoffer1998.pp199). The difficulty level of each item was determined by dividing the number of candidates who gave correct answer with the total number of sample who attempted the respective item, while the discrimination power of each item was determined by subtracting the total correct responses of the 27 percent low achievers from the total correct responses of the 27 percent high achievers. Indices of internal consistency, item difficulty, and item discrimination, obtained for each item of all the four sub-tests, are presented in tables 1-3 and annexure C.

Second Experimental Try-out

Since the results of first try out revealed the desirability of discarding some items, modifying few items and replacing some others, second try out was needed to ascertain the adequacy of the revised set of items. The second experimental tryout of proposed sub-tests (second draft) was carried out to serve the following purposes:

1. Item analysis of the sub-tests to determine the following item characteristics:

a) Internal Consistency

b) Difficulty level.

c) Discrimination power.

2. Final selection of items for each sub-test and arrangement of the items in the light of the above information.

3. Average time for proposed test and sub-tests

Sample

The second draft of proposed sub-tests was administered to a sample of 200 candidates of 105 PMA Long Course appearing at Army Selection and Recruitment Center, Rawalpindi. This sample was not only representative of

different areas of Pakistan but was essentially similar to the population for which the test was used. Random sampling procedure was used for the selection of candidates in order to have a representative sample.

Item Analysis

Item analysis was again the major objective of this experimental tryout. Each item was analyzed to find out its effectiveness, considering the role it could play in developing the final version of the test. For item analysis of the proposed sub-tests (second draft), again following three indices were calculated.

a) Internal Consistency of the items: To see the relevance of each item with the test.

b) Index of item difficulty: To see the percentages of subjects who gave correct answer to the item.

c) Index of item discrimination: To see if the test discriminate between high and low achievers on the sub-test.

To determine the internal consistency of the items and examine their relevance with the test, again item-total correlation was calculated. To determine the difficulty level and discriminatory power of the items, method of contrasted groups was used. The three indices for each item of the four sub-tests are presented in tables 4-6 and annexure E.

Time of the Test

Once the statistically effective items were selected for every sub-test, they were arranged in descending order of difficulty level, so as to get the final form of the sub-tests. In order to determine the time limit for each of the subtest, a separate study was carried out, in which all the four subtests were administered one by one to a sample of 60 candidates of the 105 PMA Long Course. The candidates were informed about the purpose of the test and were told to complete the test as quickly as possible. With the help of stopwatch, the time taken by the first candidate was noted down, followed by the time taken by the first 48 (80%) candidates out of 60, to complete the tests. The average time taken by these 48 candidates was calculated in order to decide the time limit of the test (table 13).

Main Study

After determining the time limits of the proposed test, it was decided to give the title of **Sajjad Verbal Intelligence Test in Urdu for adolescents (SVITU)** to the newly developed test. The purpose of the main study was to determine the psychometric characteristics, that is, reliability, validity, and norms of SVITU and its sub-tests.

Sample

The sample for the main study consisted of 535 candidates of 106 PMA Long Course. The sample was selected randomly from various cities of Pakistan. The sampling plan used in this study is given in Annexure-A

Administration of the Test

The procedure of the test administration in this study was the same as was used in tryout studies, except that a time limit was allocated to SVITU and its sub-test. About which the subjects were informed at the beginning of the test, and were instructed to do their best to complete each sub-test within the allotted time limit.

Item discrimination Analysis

Before determining the psychometric characteristics of SVITU and sub-tests, a study was carried out to assertain the discriminatory power of

items of the sub-tests. For this purpose, comparison of the performances of high and low groups on all items of the sub-tests was carried out. To achieve this objective, a sample of 535 candidates of 106 PMA Long Course were administered the four sub-tests of SVITU. The sample was divided into two groups designed as high scorers and low scorers on the basis of the medians of the respective sub-tests. One group consisting of high-scoring subjects falling in the upper bracket (above median) and the second low-scoring subjects falling in the lower bracket (below median) on the test. Comparisons were made on each item of the test between high and low scores in terms of their true and false responses. A 2x2 chi-square analysis was run for each item of the sub-tests (tables 14-17)

Reliability of the Test

The following methods were used to establish the reliability of SVITU and its sub-tests:-

Kuder Richardson Reliability

Split-half Reliability

Test-retest Reliability

Kuder Richardson Reliability(KR-20)

Kuder Richardson Formula 20 (KR-20) for estimating reliability is generally assumed as the best technique to find out inter-item consistency of the intelligence tests. This method is based on the numbers of the items in the test and the average inter-correlation among the test items. Responses of the items are coded as 1 for right response and 0 for incorrect response.

Test-retest Reliability

It is one of the oldest and most frequently used method of estimating the temporal stability of the test. The rationale behind this method is that since the same test is administered twice, the difference between scores on two occasions are due to measurement error. To estimate test-retest reliability, SVITU was administered to a group of 100 students with intermediate qualification and was then re-administered to the same population with a time interval of one month. The results are presented in table 18

Split-half Reliability

To determine the Split-half reliability, SVITU was administered to the sample of main study. The items of the tests were splitted into two equal halves based on odd and even numbers. Since split-half reliability is

determined by correlating the scores of two halves of the tests, an estimate of the reliability of the full test was made by applying the Spearman-Brown formula (table 19).

Validity of the Tests

When constructing a psychological test, the most important question is, to what extent will the interpretation of the scores be appropriate, meaningful and useful for the intended application of the results. Validity, one of the psychometric characteristic of the test, is the answer to this question. Validity refers to the appropriateness of the interpretations made from test scores and other evaluation results with regard to the particular use for a given group of individuals and not to the instrument itself.

Although there are many types of evidences that can be used in the process of test validation, construct validity, concurrent validity, intenal consistency, and grade differentiation were selected for the present research. Construct validity involves the demonstration of the psychological characteristics of the variable measured by the test. Concurrent validity involves prediction of the future performance and estimation of the present performance of the candidates on the test scores with reference to some valued measure as the criterion. Internal consistency was determined by correlating the scores of sub-tests with each other and with SVITU and grade

differentiation involves the differences in the performance of subjects on the tests among the four grades of the students (grades 9^{th}-13^{th}).

Construct Validity

The construct validity of a test is the extent to which the test may be said to measure a theoretical construct or trait. The goal of construct validity is to determine whether test scores provide a good measure of a specific construct. The process of construct explication provides a definition of the construct in terms of concrete behaviors. There are few methods of assessing the construct validity. The most basic method is to correlate scores on the test with scores on a number of other tests measuring the same attribute.

For this purpose of validity, criterion selected was the raw scores obtained by the candidates on Army Intelligence Test (AIT), which is currently used for selection of personnel for Armed Forces. Army Intelligence Test consists of Verbal Intelligence Test (VIT) and Non-verbal Intelligence Test (NVIT). SVITU was administered to a randomly selected group of 535 candidates of 106 PMA Long Course. Coefficients of correlation between scores (AIT & SVITU) were computed to exhibit the construct validity of SVITU and its sub-tests (table 20).

Concurrent Validity

To determine the concurrent validity of an intelligence test, the criterion frequently employed is some index of academic achievement. Criterion selected to estimate the concurrent validity of SVITU and its sub-tests was the college marks in their last annual examination, which were obtained from their educational records. However, marks of the candidates obtained in their last examination were converted into percentages for making the statistical calculations easy. Results are shown in table 21.

Internal consistency

To determine the internal consistency, the scores of all the sub-tests were correlated with each other and also with SVITU. Obtained correlation coefficients (table 22) were another measure of construct validity.

Grade differentiation

Another criterion used to determine the validity of SVITU was grade differentiation. It was hypothesized that if the newly developed test is a valid measure of intelligence, it will differentiate significantly between students varying in amount of education (years of schooling). To achieve this objective, a separate study was carried out. The SVITU was administered to a new sample of 250 students of Federal Government Educational Institutions of Lahore. This sample included 50 students from each grades 9^{th} to13^{th}. ANOVA was applied for finding the differences among the specified grades. Results are shown in the tables 23-28.

Other Statistical Analyses

To study the significance of differences between mean SVITU scores of various groups, t-test for independent groups was applied for the following comparisons:

a) male versus female students

b) English versus Urdu medium school students

c) Urban versus rural students

d) Science versus arts students

e) Sons of government servants versus businessmen

Differences in test scores of the male and female students

To determine the significance of differences between the mean SVITU scores of male and female students, t-test for independent groups was applied. The sample consisted of 200 students: 100 male students from Federal Government College Lahore and 100 female students from Government Islamia College for Girls Lahore. All these students fulfilled the predetermined criterion of education (FA/ FSc). Results of the t-test are given in table 29.

Differences in test scores of the students of the Urdu and English medium schools

To study the significance of differences between the mean SVITU scores of the students of English medium and Urdu medium schools, t-test for independent groups was applied. The sample selected for this purpose consisted of 535 candidates of PMA Long Course, having the intermediate qualification. The age range of the sample was 17-20 years. The test was administered in group settings. Results of the t-test are given in table 30.

Differences in test scores of students of the urban and rural backgrounds

To determine the significance of differences between the mean SVITU scores of the students of the urban and rural backgrounds, t-test for independent groups was applied. The sample selected for this calculation was the same of main study. Results of the t-test are given in table 31.

Differences in test scores of the students of the science and arts subjects

To see the significance of differences between the mean SVITU scores of the students of the science and arts subjects, t-test for independent groups was applied. The sample selected for this calculation was the same of main study. Results of the t-test are given in table 32.

Differences in test scores of the sons of government servants and sons of businessmen

To determine the significance of differences between the mean SVITU scores of the sons of government servants and sons of businessmen, t-test for independent groups was applied. The sample selected for this calculation was the same of main study. Results of the t-test are given in table 33.

Development of Norms

Nearly all standardized tests provide some form of within-group norms. With such norms, the individual's performance is evaluated in terms of the performance of the most nearly comparable standardization group, for instance, comparing a child's raw score with that of children of the same chronological age or in the same college grade. The key in this process is to obtain samples representing a cross section of the target population. Two different type of norms were developed for SVITU, namely Percentiles and T-scores.

Procedure

Once the final version of the test was ready for future administration, a study was conducted for developing norms of SVITU for grade 12, for Pakistani population. Development of Norms was as follows:

Sample

The sample comprised of 1080 candidates of PMA Long Courses. For developing the norms, technique of cluster sampling was used. Main step was to identify regions of the country, with reference to various demographic variables like urban-rural backgrounds, Urdu-English medium of instructions, science-arts subjects. Sampling plan is given in Annexure A.

Instrument

SVITU, comprising of 128 multiple-choice items, comprising four sub-tests (Vocabulary Test, Numerical Reasoning Test, Verbal Reasoning Test and Information Test), was administered to this sample in order to develop norms for Pakistani adolescents of grade 12.

Administration and Scoring

The test was administered to 1080 candidates of 106 PMA Long Course, belonging to the main ten cities of Pakistan as shown in the annexure-A. Total administration was completed within one month. Every group consisted of 100 candidates and they completed the test in stipulated time of 40 minutes. The students marked their responses on the answer sheets. Scoring was accomplished through computer. The maximum and minimum possible

scores on SVITU and its sub-tests are as follows:

Test	Range of scores
Vocabulary	0-42
Verbal Reasoning	0-20
Numerical Reasoning	0-36
Information	0-30
SVITU	0-128

Percentiles

Percentile scores are expressed in terms of the percentage of persons in the standardization sample who fall below a given raw score. They indicate the individual's relative position in the standardization sample. For example, if 20 percent persons give correct answers to 10 items in a vocabulary test then the raw score of 10 will correspond to the 20^{th} percentile. Percentile can be expressed as ranks (in reverse order) in a group of 100. The 50th percentile (P50) corresponds to the median. Percentile above 50 represents above average performance; those below 50 signify inferior performance. Percentiles scores are easy to compute comprehend. They are useful in explaining test results to individuals who have little background in the statistics of testing. They are universally applicable and can be used equally well with individuals of all ages, and for any type of test. However, one of the limitations of percentile scores is that they under emphasize the difference between scores

lying towards the extremes of distribution. The percentile norms of SVITU and its sub-tests are given in tables 34-38.

T Scores

T scores represent one of the linear transformations of the raw scores into the transformed scores that, in turn, act as the reference point for the interpreter to give the meaning to the obtained raw scores. T-scores refer to any set of normally distributed standard scores that has a mean of 50 and standard deviation of 10. For the calculation of T scores, first of all raw scores of SVITU and sub-tests were converted into z-scores. The obtained z-scores were then converted into T scores. Z-scores and T-scores of SVITU and its sub-tests are given in tables 39-43.

RESULTS

RESULTS

First Experimental Tryout

In evaluating the items, first consideration was given to the internal consistency of the items, calculated by item total correlation method. The difficulty levels and discrimination powers of only those items were determined that bear a significant correlation with the total score ($p<.001$). It was decided to retain the items with difficulty level ranging from .20 to .80 and with discrimination power ranging from .30 to 1.00.

Table 1

Probability coefficients showing the discriminatory power of the items of each sub-test (First Draft)

Probability Coefficient	Number of Items			
	Vocabulary	Verbal Reasoning	Numerical Reasoning	Information
.91 – 1.0	0	0	0	0
.81 - .90	0	0	0	0
.71 - .80	1	0	0	1
.61 - .70	4	0	11	6
.51 - .60	11	5	12	13
.41 - .50	21	11	15	8
.31 - .40	13	15	11	6
.21 - .30	19	7	6	3
.11 - .20	16	2	1	1
.01 - .10	10	0	0	1

Table 2

Probability coefficients showing the difficulty level of the items of each sub-test (First Draft)

Probability Coefficient	Number of Items			
	Vocabulary	Verbal Reasoning	Numerical Reasoning	Information
.91 - 1.0	3	0	0	0
.81 - .90	15	13	2	2
.71 - .80	8	18	15	3
.61 - .70	12	5	12	12
.51 - .60	15	2	11	7
.41 - .50	15	0	7	9
.31 - .40	13	1	6	4
.21 - .30	7	2	2	2
.11 - .20	6	0	1	1
.01 - .10	3	0	0	0

Vocabulary Test

Table 1 shows that discrimination power of items comprising Vocabulary Test ranged from -.05 to .72. It is evident from the table that 55 items out of 95 (57.89%) fall within the desired range (.30 and above), whereas 40 items (42.10%) fall below .30. Table 2 shows that difficulty levels of Vocabulary Test ranged from .03 to .94. As evident from table, 68 items out

of 95 (71.57%) fall within the desired range (.20 to .80). Out of the remaining 27 items, 18 were too easy (p=.80 to .94) and 9 items were too difficult (p=.03 to .19) and hence were discarded. Results reveal that 42 items (44.21%) fall within the range of .41 to .70 of difficulty level while 36 items (37.89%) fall within the range of .41 to .70 of discrimination power.

Verbal Reasoning Test

Table 1 shows that discrimination power of Verbal Reasoning Test ranged from .20 to .55. Table further reveals that 31 items out of 40 (77.50%) fall within the desired range (.30 and above), whereas 9 items (22.50%) fall below .30. Table 2 shows that difficulty levels of Verbal Reasoning Test ranged from .23 to .89, Results show that 24 items out of 40 (60%) fall within the desired range (.20 to .80), 16 items were too easy (p=.81 to .89), and none of the items was below .23 level of difficulty. Results reveal that 7 items (17.50%) fall in the range of .41 to .70 of difficulty level while 16 items (40%) fall within the range of .41 to .70 of discrimination power.

Numerical Reasoning Test

Table 1 shows that discrimination power of items comprising Numerical Reasoning Test ranged from .16 to .70. It is evident from the table that 51 items out of 56 (91.07%) fall within the desired range (.30 and above),

whereas 5 items (8.92%) fall below .30. Table 2 shows that difficulty level of Numerical Reasoning Test ranged from .18 to .83. As is evident from table, 52 items out of 56 (94.64%) fall within the desired range (.20 to .80). Out of the remaining 4 items, 3 were too easy (p=.81 to .83) and one item was too difficult (p=.19) and hence were discarded. Results reveal that 30 items (53.57%) fall in the range of .41 to .70 of difficulty level while 38 items (67.85%) fall within the range of .41 to .70 of discrimination power.

Information Test

Table 1 shows that discrimination power of items comprising Information Test ranged from .08 to .70. It is evident from the table that 36 items out of 40 (90%) fall within the desired range (.30 and above), whereas 4 items (10%) fall below .30. Table 2 shows that difficulty level of Information Test ranged from .03 to .94. As evident from table, 37 items out of 40 (92.50%) fall within the desired range (.20 to .80). Out of the remaining 3 items, 2 were too easy (p=.81 to .84) and one item was too difficult (p=.08-.19) and hence were discarded. Results reveal that 28 items (70%) fall within the range of .41 to .70 of difficulty level while 27 items (67.50%) fall within the range of .41 to .70 of discrimination power.

Table 3

The number of items discarded from the test during the first experimental try-out

	Items		
Sub-test	Total	Discarded*	Selected
Vocabulary	95	41	54
Verbal Reasoning	40	30	10
Numerical Reasoning	56	18	38
Information	40	4	36

*Vocabulary =1, 3, 6, 8, 9, 12, 13, 16, 19, 20, 21, 23, 24, 25, 30, 34, 35, 36, 38, 39, 42, 45, 47, 50, 51, 52, 53, 56, 58, 59, 61, 63, 64, 66, 68, 69, 76, 86, 90, 92, 95.

*Verbal Reasoning =1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 19, 22, 24, 25, 26, 27, 28, 31, 32, 34, 37, 38, 40.

*Numerical Reasoning = 3, 4, 5, 10, 16, 17, 22, 24, 27, 33, 37, 38, 40, 41, 43, 51, 54, 56.

*Information = 2, 3, 33, 38.

On the basis of the selection criteria for the effective items, all those items, which proved ineffective, obscure or complex, were discarded from the item pool as shown in the table 3. As evident from table, the percentages of items discarded from Vocabulary Test, Verbal Reasoning Test, Numerical Reasoning Test, and Information Test were 43 percent, 75 percent, 32 percent, and 10 percent, respectively.

Second Experimental Tryout

In evaluating the items, again first consideration was given to the internal consistency of the items, calculated by item total correlation method. The difficulty levels and discrimination powers of only those items were determined that bear a significant correlation with the total score ($p<.001$). It was decided to retain the items with difficulty level ranging from .20 to .80 and with discrimination power ranging from .30 to 1.00.

Table 4

Probability coefficients showing the discriminatory power of the items of each sub-test (Second Draft)

Probability Coefficient	Number of Items			
	Vocabulary	Verbal Reasoning	Numerical Reasoning	Information
.91 – 1.0	0	1	0	0
.81 - .90	0	0	0	0
.71 - .80	3	1	0	1
.61 - .70	5	0	3	3
.51 - .60	3	0	6	9
.41 - .50	11	1	9	8
.31 - .40	10	4	12	9
.21 - .30	12	11	7	1
.11 - .20	5	4	2	5
.01 - .10	2	6	3	0

Table 5

Probability coefficients showing the difficulty levels of the items of each sub-test (Second Draft)

Probability Coefficient	Number of Items			
	Vocabulary	Verbal Reasoning	Numerical Reasoning	Information
.91 - 1.0	2	3	1	2
.81 - .90	4	8	9	2
.71 - .80	6	4	9	10
.61 - .70	15	3	9	11
.51 - .60	10	0	4	3
.41 - .50	5	5	5	4
.31 - .40	7	2	3	3
.21 - .30	5	2	0	1
.11 - .20	1	1	0	1
.01 - .10	0	0	2	1

Vocabulary Test

Table 4 shows that discrimination power of items comprising Vocabulary Test ranged from .01 to .80. It is evident from the table that 37 items out of 54 (68.51%) fall within the desired range (.30 and above), whereas 17 items (31.48%) fall below .30. Table 5 shows that difficulty levels of Vocabulary Test ranged from .14 to .96. As evident from table, 46 items out of 54 (85.18%) fall within the desired range (.20 to .80). Out of the remaining 8 items, 6 were too easy (p =.81 to .96) and 2 items were too difficult (p =.14 to.19) and hence were discarded. Results reveal that 30 items (55.55%) fall within the range of .41 to .70 of difficulty level while 19 items (35.18%) fall within the range of .41 to .70 of discrimination power.

Verbal Reasoning Test

Table 4 shows that discrimination power of Verbal Reasoning Test ranged from 0 to .61. Table reveals that 10 items out of 33 (30.30%) fall within the desired range (.30 and above), whereas 23 items (69.69%) fall below .30. Table 5 shows that difficulty level of Verbal Reasoning Test ranged from .22 to 1.00. Results show that 9 items out of 33 (27.27%) fall within the desired range (.20 to .80). The remaining 24 items were too easy (p=.81 to .89), and none of the items was below .23 level of difficulty. Results reveal that 8 items (24.24%) fall within the range of .41 to .70 of difficulty level

while only one item (3.03) fall within the range of .41 to .70 of discrimination power.

Numerical Reasoning Test

Table 4 shows that discrimination power of items comprising Numerical Reasoning Test ranged from 0 to .67. It is evident from the table that 30 items out of 42 (71.42%) fall within the desired range (.30 and above), whereas 12 items (28.57%) fall below .30. Table 2 shows that difficulty level of Numerical Reasoning Test ranged from 0 to .92. As is evident from table, 31 items out of 42 (73.80%) fall within the desired range (.20 to .80). Out of the remaining 11 items, 9 were too easy (p=.81 to .92) and 2 items were too difficult (p=.19 and below) and hence were discarded. Results reveal that 18 items (42.85%) fall within the range of .41 to .70 of difficulty level while 18 items (42.85%) fall within the range of .41 to .70 of discrimination power.

Information Test

Table 4 shows that discrimination power of items comprising Information Test ranged from .11 to .72. It is evident from the table that 30 items out of 36 (83.33%) fall within the desired range (.30 and above), whereas 6 items (16.66%) fall below .30. Table 5 shows that difficulty level of Information Test ranged from .26 to .92. As evident from table, 32 items out

of 36 (88.88%) fall within the desired range (.20 to .80). The remaining 4 items were too easy (p=. 81 to .92), hence were discarded. Results reveal that 18 items (50%) fall within the range of .41 to .70 of difficulty level while 20 items (55.55%) fall within the range of .41 to .70 of discrimination power.

Tables 4-5 show that in Vocabulary Test, 48 items fall within .20-.80 level of difficulty, while 44 items fall within .20-.80 level of discrimination. In Verbal Reasoning Test, 16 items fall within .20-.80 level of difficulty, while 17 items fall within .20-.80 level of discrimination. In Numerical Reasoning Test, 30 items fall within .20 -.80 level of difficulty, while 37 items fall within .20-.80 level of discrimination. Whereas in Information Test, 32 items fall within .20 -.80 level of difficulty, while 31 items fall within .20-.80 level of discrimination. It was again decided to keep the items with difficulty level ranging from .20 to .80 with discrimination power ranging from .30 to 1.00.

Table 6

The number of items discarded from the test during the second experimental try-out

	Items		
Sub-test	Total	Discarded*	Selected
Vocabulary	54	12	42
Verbal Reasoning	29	5	24
Numerical Reasoning	45	9	36
Information	36	6	30

*Vocabulary = 1,6,7,8,10,13,20,24,25,31,34,42

*Verbal Reasoning = 7, 10,17,23,28.

*Numerical Reasoning = 2,7,11,14,23,25,27,35,42

*Information = 1,2,3,4,9

On the basis of the selection criteria for the effective items, all those items, which proved ineffective, obscure or complex, were discarded from the item pool as shown in the table 6. As evident from table, the percentages of items discarded from Vocabulary Test, Verbal Reasoning Test, Numerical Reasoning Test, and Information Test were 22 percent, 17 percent, 20 percent, and 14 percent, respectively.

Table 6 shows that 54 items of Vocabulary Test were reduced to 42, 33 items of Verbal Reasoning Test were reduced to 30, 45 items of Numerical Reasoning Test were reduced to 36, and 36 items of Information Test were reduced to 30 items. All the selected items were rearranged in each sub-test according to their descending order of difficulty level (tables 7-10) and their dispersion with reference to difficulty level and discrimination power are shown in Figures 1 to 4 as shown in Annexure-G.

Table 7

Three psychometric indices of the items of Vocabulary Test (Final Draft)

Items	Difficulty level	Discrimination level	Internal Consistency
1.	82.41	.28	.42
2.	80.56	.31	.34
3.	75.93	.33	.33
4.	75.00	.39	.38
5.	75.00	.39	.36
6.	74.07	.41	.34
7.	73.15	.46	.43
8.	68.52	.33	.30
9.	66.67	.37	.28
10.	66.67	.63	.47
11.	65.74	.28	.25
12.	65.74	.43	.39
13.	65.74	.24	.20
14	64.81	.56	.44
15	64.81	.44	.38
16.	64.81	.63	.52
17.	63.89	.24	.23
18.	62.96	.19	.20
19.	62.96	.56	.46
20.	62.96	.44	.36
21.	62.04	.39	.36
22.	59.26	.67	.51
23.	55.56	.70	.54
24.	55.56	.30	.24
25.	54.63	.57	.45
26	52.78	.72	.50
27.	51.85	.30	.24
28.	46.30	.41	.35

Cont.....

29.	46.30	.33	.32
30.	45.37	.28	.25
31.	44.44	.59	.46
32.	43.52	.76	.61
33	39.81	.69	.55
34	37.96	.31	.32
35.	33.33	.48	.41
36.	32.41	.28	.26
37.	31.48	.44	.44
38.	31.48	.26	.25
39.	29.63	.33	.34
40.	28.70	.28	.21
41.	21.30	.20	.25
42.	20.37	.19	.20

Table 8

Three psychometric indices of the items of Verbal Reasoning Test(Final Draft)

Items	Difficulty level	Discrimination level	Internal Consistency
1.	89.81	.20	.50
2.	89.81	.20	.57
3.	89.81	.20	.60
4.	89.81	.20	.51
5.	89.81	.20	.52
6.	87.96	.20	.56
7.	87.04	.26	.59
8.	86.11	.24	.39
9.	86.11	.20	.29
10.	84.26	.31	.44
11.	81.48	.37	.34
12.	80.56	.35	.45
13.	79.63	.33	.55
14.	76.85	.43	.42
15.	75.00	.46	.40
16.	75.00	.31	.29
17.	64.81	.37	.36
18.	63.89	.61	.51
19.	60.19	.39	.33
20.	22.22	.26	.20

Table 9

Three psychometric indices of the items of Numerical Reasoning Test (Final Draft)

Items	Difficulty level	Discrimination level	Internal Consistency
1.	86.11	.28	.49
2.	86.11	.24	.48
3.	85.19	.30	.54
4.	85.19	.26	.53
5.	85.19	.19	.42
6.	81.48	.37	.57
7.	79.63	.33	.44
8.	78.70	.28	.38
9.	77.78	.37	.61
10.	76.85	.43	.50
11.	76.85	.39	.48
12.	75.00	.46	.49
13.	75.00	.31	.44
14.	73.15	.46	.52
15.	71.30	.35	.41
16.	69.44	.39	.42
17.	67.59	.50	.50
18.	67.59	.35	.40
19.	65.74	.35	.33
20.	64.81	.63	.55
21.	64.81	.52	.41
22.	62.96	.48	.42
23.	61.11	.44	.41
24.	61.11	.44	.36
25.	56.48	.46	.39
26.	54.63	.65	.52

27.	51.85	.52	.40
28.	50.93	.54	.42
29.	47.22.	.57	.39
30.	47.22	.39	.29
31.	43.52	.54	.38
32.	43.52	.46	.34
33.	42.59	.67	.43
34.	36.11	.35	.29
35.	35.70	.44	.37
36.	34.26	.24	.18

Table 10

Three psychometric indices of the items of Information Test (Final Draft)

Items	Difficulty level	Discrimination level	Internal Consistency
1.	84.26	.31	.67
2.	81.48	.37	.70
3.	78.70	.35	.54
4.	77.78	.44	.64
5.	75.93	.48	.66
6.	74.07	.48	.56
7.	73.15	.50	.63
8.	73.15	.31	.49
9.	72.22	.56	.70
10.	71.30	.50	.58
11.	70.37	.52	.54
12.	70.37	.41	.52
13.	67.59	.46	.51
14.	66.67	.56	.44
15.	66.67	.52	.52
16.	66.67	.37	.35
17.	65.74	.39	.39
18.	64.81	.44	.52
19.	63.89	.72	.60
20.	62.96	.63	.53
21.	62.04	.61	.48
22.	61.11	.41	.36
23.	60.19	.57	.51
24.	58.33	.50	.40
25.	54.63	.61	.51
26.	53.70	.56	.47
27.	49.07	.39	.30
28.	48.15	.48	.42
29.	38.89	.22	.25
30.	32.41	.39	.29

Table 11

Probability coefficients showing the discriminatory power of the items of each sub-test(Final draft)

Probability Coefficient	Number of Items			
	Vocabulary	Verbal Reasoning	Numerical Reasoning	Information
.91 - 1.0	0	0	0	0
.81 - .90	0	0	0	0
.71 - .80	2	0	0	1
.61 - .70	5	1	3	3
.51 - .60	4	0	5	6
.41 - .50	8	2	10	11
.31 - .40	10	7	11	8
.21 - .30	10	4	6	1
.11 - .20	3	7	1	0
.01 - .10	0	0	0	0

Table 12

Probability coefficients showing the difficulty levels of the items of each sub-test(Final draft)

Probability Coefficient	Number of Items			
	Vocabulary	Verbal Reasoning	Numerical Reasoning	Information
.91 – 1.0	0	0	0	0
.81 - .90	2	11	6	2
.71 - .80	5	5	9	8
.61 - .70	14	3	9	12
.51 - .60	6	0	4	4
.41 - .50	5	0	5	2
.31 - .40	6	0	3	2
.21 - .30	4	1	0	0
.11 - .20	0	0	0	0
.01 - .10	0	0	0	0

Average time needed for SVITU and sub-tests

The time limits for newly developed test, hereafter called Sajjad Verbal Intelligence Test for adolescents (SVITU) and its sub-tests were determined by computing the average time taken by 80% of the subjects who completed each sub-test. As shown in table 13, the maximum time required by majority of subjects is 40 minutes which was taken as the time limit for the newly developed test SVITU

Table 13

Average time needed for SVITU and its sub-tests

Tests	Items	Time in minutes
Vocabulary	42	16
Verbal Reasoning	20	6
Numerical Reasoning	36	12
Information	30	6
SVITU	128	40

Item discrimination Analysis

Chi-square test was applied to each item of the sub-tests. The Chi-square values of all the items of the sub-tests are listed in the tables 14-17. These indices of the items also serve as the evidence of item validity.

Table 14

Chi-square analysis of true and false responses of high and low scoreres on each item of Vocabulary Test

Items	Low Group (n=270) No. of Responses		High Group (n=265) No. of Responses		Chi-Square Value
	True	False	True	False	
1.	216	54	354	12	29.60**
2.	111	159	181	84	39.88**
3.	215	55	249	16	23.86**
4.	193	77	253	12	55.50**
5.	185	85	233	32	29.47**
6.	161	109	249	16	88.04**
7.	23	247	38	22	4.48*
8.	143	127	198	67	27.38**
9.	154	116	243	23	81.75**
10.	111	159	190	75	50.84**
11.	142	128	210	55	42.21**
12.	198	72	215	50	4.62*
13.	100	170	183	82	50.03**
14.	144	126	211	54	41.40**
15.	72	198	188	77	104.95**
16.	52	216	197	68	163.07**
17.	164	106	209	56	20.81**
18.	143	127	234	31	80.25**
19.	143	127	232	33	76.30**
20.	103	167	213	52	98.64**
21.	94	177	198	67	87.43**
22.	104	166	186	79	54.03**
23.	75	195	187	78	97.98**
24.	127	143	206	59	53.63**
25.	112	158	235	30	130.77**
26.	81	189	157	108	46.31**
27.	65	205	171	94	88.77**

Cont...

28.	91	179	221	44	135.86**
29.	81	189	214	51	139.37**
30.	86	184	160	105	43.81**
31.	102	168	176	89	43.93**
32.	64	206	153	112	64.24**
33.	56	214	129	136	46.14**
34.	52	218	124	141	45.92**
35.	40	230	163	102	123.84**
36.	48	222	164	101	108.76**
37.	51	219	97	168	20.97**
38.	38	232	131	134	77.37**
39.	37	233	115	150	57.97**
40.	54	216	127	138	46.58**
41.	50	220	107	158	30.82**
42.	44	226	101	164	32.21**

$*p<.01$ $**p<.0001$

Table 15

Chi-square analysis of true and false responses of high and low scoreres on each item of Verbal Reasoning Test

Items	Low Group (n=270) No. of Responses		High Group (n=265) No. of Responses		Chi-Square Value
1.	316	27	192	0	15.91**
2.	338	5	191	1	.97
3.	339	4	191	1	.55
4.	306	37	189	3	15.14**
5.	301	42	191	1	22.89**
6.	282	61	186	6	24.14**
7.	319	24	190	2	9.44*
8.	311	32	191	1	16.50**
9.	276	67	181	1	18.83**
10.	288	55	188	4	24.42**
11.	319	24	192	0	14.06**
12.	295	48	191	1	26.85**
13.	261	82	189	3	45.98**
14.	244	99	184	8	46.92**
15.	195	148	180	1	79.95**
16.	319	24	191	1	11.59*
17.	114	229	135	5	68.01**
18.	205	138	172	20	52.58**
19.	123	220	150	42	87.99**
20.	61	282	86	106	45.05**

$*p<.01$ $**p<.0001$

Table 16

Chi-square analysis of true and false responses of high and low scoreres on each item of Numerical Reasoning Test

Items	Low Group (n=270)		High Group (n=265)		Chi-Square Value
	No. of Responses		No. of Responses		
	True	False	True	False	
1.	266	29	236	4	15.23**
2.	247	48	230	10	20.06**
3.	254	41	228	12	11.74**
4.	279	16	236	4	5.19*
5.	278	17	238	2	9.38**
6.	266	29	235	5	13.34**
7.	222	73	219	21	23.37**
8.	260	35	238	2	25.01**
9.	267	28	235	5	12.54**
10.	242	53	223	17	13.78**
11.	203	92	219	21	39.98**
12.	208	87	223	17	42.43**
13.	209	86	224	16	43.36**
14.	203	92	206	34	21.29**
15.	181	114	200	40	31.18**
16.	193	102	222	18	55.76**
17.	181	114	212	28	49.39**
18.	130	165	150	90	18.02**
19.	168	127	206	34	52.48**
20.	150	145	204	36	68.95**
21.	117	178	178	62	63.69**
22.	115	180	172	68	56.84**
23.	141	154	195	45	63.39**
24.	92	203	191	49	124.40**
25.	132	163	193	47	70.61**
26.	74	221	159	81	91.21**
27.	106	189	176	64	74.26**
28.	172	123	210	30	55.23**
29.	67	228	150	90	86.89**
30.	51	244	116	124	59.40**
31.	145	150	219	21	107.84**
32.	58	237	173	67	148.22**
33.	64	231	124	116	52.50**
34.	104	191	197	43	117.93**
35.	18	277	77	163	61.17**
36.	22	273	82	158	60.28**

*$p<.01$ **$p<.0001$*

Table 17

Chi-square analysis of true and false responses of high and low scoreres on each item of Information Test

Items	Low Group (n=270) No. of Responses		High Group (n=265) No. of Responses		Chi-Square Value
	True	False	True	False	
1.	`250	25	259	1	21.91**
2.	249	26	260	0	25.83**
3.	202	73	246	14	43.94**
4.	230	45	254	6	30.61**
5.	203	72	239	21	30.50**
6.	205	70	253	07	56.20**
7.	187	88	251	9	73.32**
8.	210	65	251	9	45.64**
9.	199	76	255	5	68.77**
10.	159	116	237	23	77.22**
11.	158	117	237	23	78.55**
12.	183	92	254	6	86.65**
13.	192	83	245	15	53.23**
14.	120	155	223	37	103.11**
15.	127	148	215	45	77.25**
16.	121	154	208	52	73.14**
17.	131	144	222	38	84.84**
18.	164	111	243	17	84.01**
19.	155	120	243	17	96.55**
20.	46	229	70	190	8.81*
21.	125	150	199	61	54.06**
22.	117	158	221	39	103.54**
23.	137	138	238	22	110.96**
24.	131	144	218	42	77.26**
25.	120	155	207	53	72.80**
26.	92	183	187	73	79.25**
27.	59	216	119	150	26.89**
28.	67	208	177	83	102.94**
29.	64	211	118	142	29.11**
30.	61	214	134	126	49.72**

*$p<.01$ **$p<.0001$

Reliability of the Test

As described earlier (pp 71-73), the reliability of the newly developed test SVITU and sub-tests was determined by the following methods:

a) Kuder Richardson Reliability

b) Split-half Reliability

c) Test -retest Reliability

Table 18

Reliability coefficients of SVITU and its sub-tests

Test	KR-20	Test-retest
Vocabulary	.87*	.80*
Verbal Reasoning	.70*	.68*
Numerical Reasoning	.85*	.81*
Information	.90*	.83*
SVITU	.92*	.86*

*P<.0001

Table 18 shows that all reliability coefficients obtained by Kuder-Richardson method are high and as such demonstrate well the internal consistency of SVITU and all the four sub-tests. These indices further indicate that items of SVITU and its sub-tests are highly homogenous. These results suggest that the newly developed test is a reliable instrument for measuring intelligence of adolescents of grade 12 in Pakistan. The test-retest reliability coefficients of SVITU and its sub-tests (table 18) are all significant statistically.

Table 19

Split-half Reliability of SVITU and its sub-tests

Test	Odd-even	After correction for length
Vocabulary	.78*	.87*
Verbal Reasoning	.52*	.69*
Numerical Reasoning	.76*	.86*
Information	.84*	.90*
SVITU	.86*	.92*

*$p<.0001$

The spilt-half reliability coefficients of SVITU and its sub-tests (table 19) are all significant statistically. After applying Spearman-Brown correction formula for test length, the indices seem to be quite high for total length of SVITU and its sub-tests.

Validity of the Tests

The following types of validity were computed for SVITU and its sub-tests.

Construct Validity

Table 20

Construct Validity of SVITU and sub-tests

Tests	Army Intelligence Test		
	Verbal	Nonverbal	Total
Vocabulary	.31*	.10*	.24*
Verbal Reasoning	.32*	.26*	.30*
Numerical Reasoning	.46*	.37*	.42*
Information	.27*	.18*	.23*
SVITU	.47*	.30*	.40*

*p< .0001

Coefficients of correlation between scores of Ss on SVITU and AIT, presented in table 20, are quite high and demonstrate well the construct validity of the newly developed test.

Concurrent Validity

Table 21

Concurrent Validity of SVITU and sub-tests

Tests	College Marks
Vocabulary	.21*
Verbal Reasoning	.18*
Numerical Reasoning	.23*
Information	.16*
SVITU	.27*

*p< .0001

Statistically significant correlations between the scores of Ss on SVITU and marks obtained in the last examination (table 21) are an evidence of the concurrent validity of the test.

Internal consistency of the sub-tests

Table 22

Internal consistency of SVITU and the subtests

Tests	Vocabulary	Verbal Reasoning	Numerical Reasoning	Information	SVITU
Vocabulary	1.00				
Verbal Reasoning	.29*	1.00			
Numerical Reasoning	.33*	.39*	1.00		
Information	.32*	.18*	.31*	1.00	
SVITU	.78*	.51*	.72*	.70*	1.00

*p< .0001

Results (Table 22) show correlation matrix, revealing the internal consistency of SVITU and its sub-tests. Statistically significant results also exhibit the evidence of construct validity of the test.

One way ANOVA along with the means and standard deviations of the students of grades 9th to 13th are shown in tables 23-28.

Table 23

Means and Standard Deviations of Vocabulary Test for Grades 9th to 13th

Grade	Mean	Standard Deviation
9th	14.96	4.74
10th	14.12	5.04
11th	23.34	5.32
12th	20.02	7.74
13th	24.42	7.58

N=50

Table 24

Means and Standard Deviations of Verbal Reasoning Test for Grades 9th to 13th

Grade	Mean	Standard Deviation
9th	4.88	1.82
10th	5.14	1.82
11th	5.38	1.49
12th	5.08	1.48
13th	5.40	1.30

N=50

Table 25

Means and Standard Deviations of Numerical Reasoning test for Grades 9th to 13th

Grade	Mean	Standard Deviation
9th	19.58	7.29
10th	17.12	6.29
11th	21.08	11.30
12th	22.38	9.55
13th	25.88	6.72

N=50

Table 26

Means and Standard Deviations of Information Test for Grades 9th to 13th

Grade	Mean	Standard Deviation
9th	18.48	6.47
10th	17.48	5.07
11th	18.56	9.39
12th	20.84	6.25
13th	22.22	5.66

N=50

Table 27

Means and Standard Deviations of SVITU for Grades 9th to 13th

Grade	Mean	Standard Deviation
9th	57.90	14.33
10th	53.86	11.64
11th	68.36	19.96
12th	68.32	18.00
13th	77.92	15.79

N=50

Table 28

One way Analysis of variance showing grade differentiation (Grades 9^{th}- 13^{th}) in terms of scores on SVITU and its sub-tests

Tests	Mean	Standard Deviation	F value
Vocabulary	19.37	7.48	28.61**
Verbal Reasoning	5.17	1.60	.925
Numerical Reasoning	21.20	8.84	7.51**
Information	19.51	6.91	4.18*
SVITU	65.27	18.20	17.29**

*p<.001**p<.0001

Differences in test scores of the male and female students

Table 29

Mean, Standard Deviation, and t-values showing significance of gender differences on SVITU and its sub-tests

Tests	Male (n=100)		Female (n=100)		t-Value
	M	SD	M	SD	
Vocabulary	23.82	6.72	24.76	7.57	.928
Verbal Reasoning	5.23	1.68	5.69	1.46	2.-06
Numerical Reasoning	23.22	5.17	23.95	6.21	.903
Information	22.60	3.67	21.42	6.05	1.66*
SVITU	74.87	12.64	75.82	13.79	.508

*df= 198 *p<.01*

To determine if SVITU scores show any gender difference, t-test for independent groups was applied to the mean SVITU scores of male and female subjects. Results (table 29) show no significant difference between male and female students, except on Information Test.

Differences in test scores of the students of Urdu and English medium schools

Table 30

Mean, Standard Deviation, and t-values showing significance of differences between students of English and Urdu medium schools on SVITU and its sub-tests

Tests	English Medium (n= 94)		Urdu medium (n= 438)		t-Value
	M	SD	M	SD	
Vocabulary	19.04	8.80	23.22	7.49	4.76**
Verbal Reasoning	16.68	2.16	16.34	2.54	1.17
Numerical Reasoning	24.30	6.29	24.06	5.95	.36
Information	19.70	6.27	20.65	6.45	1.30
SVITU	79.73	17.32	84.29	15.80	.013*

*df=530 *p<.01**p<.0001*

It was also explored if medium of instruction exerts any influence on SVITU. Our data shows (table 30) that significant differences in scores exists between English medium and Urdu medium students on Vocabulary Test and SVITU, whereas no significant difference between the two groups on Verbal Reasoning Test, Numerical Reasoning Test, and Information Test.

Differences in test scores of the students of urban and rural backgrounds

Table 31

Mean, Standard Deviation, and t-values showing significance of differences between students of urban and rural backgrounds on SVITU and its sub-tests

Tests	Urban (n= 280)		Rural (n= 253)		t-Value
	M	SD	M	SD	
Vocabulary	21.91	8.41	23.15	7.24	1.81*
Verbal Reasoning	16.60	2.54	16.19	2.40	1.91*
Numerical Reasoning	24.52	6.06	23.66	5.91	1.64
Information	20.67	5.86	20.30	7.00	.67
SVITU	83.71	16.14	83.31	16.19	.29

df=531 *p<. 01

To determine if SVITU scores show any residential difference, t-test for independent groups was applied to the mean scores of students of urban and rural backgrounds. Table 31 reveals that significant differences between students of urban and rural backgrounds were found in Vocabulary Test and Verbal Reasoning Test, whereas there was no significant difference on Numerical Reasoning Test, Information Test, and SVITU.

Differences in test scores of the students of the science and arts subjects

Table 32

Mean, Standard Deviation, and t-values showing significance of differences between students of Science and Arts on SVITU and its sub-tests

Tests	Science (n=321)		Arts (n= 201)		t-Value
	M	SD	M	SD	
Vocabulary	23.11	8.02	21.46	7.65	2.32*
Verbal Reasoning	16.69	2.34	15.92	2.62	3.49**
Numerical Reasoning	24.82	6.16	22.80	5.67	3.76***
Information	21.21	6.00	19.49	6.90	3.02**
SVITU	85.84	16.31	79.68	15.66	4.26***

*df=520 *p<.01**p<.001***p<.0001*

To determine if SVITU scores show any subjects difference, t-test for independent groups was applied to the mean scores of science and arts students. Table 32 shows that significant differences were found in SVITU and its sub-tests.

Differences in test scores of the sons of government servants and sons of businessmen

Table 33

Mean, Standard Deviation, and t-values showing significance of differences between sons of government servants and sons of businessmen on the SVITU and its sub-tests

Tests	Sons of Government Servants (n= 199)		Sons of Businessmen (n= 294)		t-value
	M	SD	M	SD	
Vocabulary	22.33	7.64	22.43	7.97	.14
Verbal Reasoning	16.60	2.50	16.19	2.52	1.76*
Numerical Reasoning	24.67	5.82	23.72	5.87	1.76*
Information	21.24	5.69	19.79	6.92	2.45*
SVITU	84.85	15.20	82.15	16.39	1.85*

df=491 *p<. 01

To determine if SVITU scores show any difference in father's occupation, t-test for independent groups were applied to the mean scores of male and female subjects. Table 33 reveals that significant difference between sons of government servants and sons of businessmen were found in Verbal Reasoning Test, Numerical Reasoning Test, Information Test and SVITU, whereas there was no significant difference on Vocabulary Test.

Norms Development

Two types of norms were computed for the normative sample of 1000 candidates representing ten main cities of Pakistan, the detail of which is given in Annexure A.

Percentiles

The maximum attainable score on the Vocabulary test was 42 whereas the minimum possible score was zero. The highest score obtained by the students on the test was 40 while the lowest score was zero The maximum attainable score on the Verbal Reasoning Test was 20 whereas the minimum possible score was zero. The highest score obtained by the students on the total test was 20 while the lowest score was zero. The maximum attainable score on the Numerical ability test was 36 whereas the minimum possible score was zero. The highest score obtained by the students on the test was 35 while the lowest score was zero. The maximum attainable score on the Information Test was 30 whereas the minimum possible score was zero. The highest score obtained by the students on the total test was 29 while the lowest score was zero. The maximum attainable score on the total test was 128 whereas the minimum possible score was zero. The highest score obtained by the students on the total test was 122 while the lowest score was zero. Scores in

of Ss normative sample of the present study were converted into percentiles. Separate percentiles were computed for total test scores and for the scores of each sub-test. Percentile norms of SVITU and its sub-tests are shown in tables 34-38.

Table 34

Percentile scores of Vocabulary Test (N = 1080)

Percentiles	Raw Scores	Percentiles	Raw Scores
1.00	3	55.00	23
5.00	8	60.00	24
10.00	11	65.00	25
15.00	12	70.00	26
20.00	14	75.00	28
25.00	15	80.00	29
30.00	17	85.00	31
35.00	18	90.00	33
40.00	19	95.00	35
45.00	20	99.00	38
50.00	21		

Table 35

Percentile scores of Verbal Reasoning Test (N = 1080)

Percentiles	Raw Scores	Percentiles	Raw Scores
1.00	5.8	55.00	17
5.00	11	60.00	17
10.00	13	65.00	17
15.00	14	70.00	18
20.00	14	75.00	18
25.00	15	80.00	18
30.00	15	85.00	18
35.00	16	90.00	19
40.00	16	95.00	19
45.00	16	99.00	20
50.00	17		

Table 36

Percentile scores for Numerical Reasoning Test (N = 1080)

Percentiles	Raw Scores	Percentiles	Raw Scores
1.00	3	55.00	24
5.00	11	60.00	25
10.00	14	65.00	26
15.00	16	70.00	27
20.00	18	75.00	28
25.00	19	80.00	29
30.00	20	85.00	30
35.00	21	90.00	31
40.00	22	95.00	32
45.00	23	99.00	34
50.00	24		

Table 37

Percentile Norms for Information Test (N = 1080)

Percentiles	Raw Scores	Percentiles	Raw Scores
1.00	0	55.00	22
5.00	0	60.00	23
10.00	9	65.00	23
15.00	13	70.00	24
20.00	16	75.00	24
25.00	18	80.00	25
30.00	19	85.00	25
35.00	20	90.00	26
40.00	20	95.00	27
45.00	21	99.00	28
50.00	22		

Table 38

Percentile Norms for SVITU (N = 1080)

Percentiles	Raw Scores	Percentile	Raw Score
1.00	34	55.00	83
5.00	49	60.00	85
10.00	57	65.00	88
15.00	62	70.00	90
20.00	65	75.00	93
25.00	69	80.00	96
30.00	72	85.00	99
35.00	74	90.00	102
40.00	76	95.00	106
45.00	78	99.00	114
50.00	81		

Linear Transformations

Z scores and T scores of SVITU and its sub-tests are shown in tables 39-42.

Table 39

Z-Scores and T-Scores equivalents of the raw scores for Vocabulary Test (N = 1080)

Scores	Z-Scores	T-Scores	Raw Scores	Z-Score	T-Scores
1	-2.45	25	22	.07	51
2	-2.33	27	23	.19	52
3	-2.21	28	24	.31	53
4	-2.09	29	25	.43	54
5	-1.97	30	26	.55	55
6	-1.85	31	27	.67	57
7	-1.73	32	28	.79	58
8	-1.61	34	29	.91	59
9	-1.49	35	30	1.02	60
10	-1.37	36	31	1.14	61
11	-1.25	37	32	1.26	63
12	-1.13	39	33	1.38	64
13	-1.01	40	34	1.50	65
14	-.89	41	35	1.62	66
15	-.77	42	36	1.75	67
16	-.65	43	37	1.87	69
17	-.53	45	38	1.99	70
18	-.41	46	39	2.10	71
19	-.29	47	40	2.23	72
20	-.17	48	41	2.35	73
21	-.05	49	42	2.47	75

M=21.43; SD=8.32

Table 40

Z-Scores and T-Scores equivalents of the raw scores for Verbal Reasoning Test (N = 1080)

Raw Scores	Z-Scores	T-Scores	Raw Scores	Z-Scores	T-Scores
1	-5.50	0	11	-1.83	32
2	-5.13	0	12	-1.47	35
3	-4.76	2	13	-1.10	39
4	-4.39	6	14	-.73	43
5	-4.03	10	15	-.36	46
6	-3.66	13	16	-.002	50
7	-3.29	17	17	.36	54
8	-2.93	21	18	.73	57
9	-2.56	24	19	1.09	61
10	-2.20	28	20	1.46	65

M=16.00; SD=2.73

Table 41

Z-Score and T-Scores Equivalents of the raw scores for Numerical Reasoning Test (N = 1080)

Raw Scores	Z-Scores	T-Scores	Raw Scores	Z-Scores	T-Score
1	-3.33	17	19	-.59	44
2	-3.19	18	20	-.44	45
3	-3.03	20	21	-.29	47
4	-2.88	21	22	-.14	48
5	-2.72	23	23	.01	50
6	-2.57	24	24	.16.	52
7	-2.42	26	25	.31	53
8	-2.27	27	26	.47	55
9	-2.12	29	27	.62	56
10	-1.96	30	28	.77	58
11	-1.81	32	29	.92	59
12	-1.66	33	30	1.07	61
13	-1.51	35	31	1.22	62
14	-1.36	36	32	1.38	64
15	-1.21	38	33	1.53	65
16	-1.05	39	34	1.68	67
17	-.90	41	35	1.83	68
18	-.75	42	36	1.98	70

M=22.93; SD=6.56

Table 42

Z-Scores and T-Scores Equivalents of the raw scores for Information Test (N = 1080)

Raw Scores	Z-Scores	T-Scores	Raw Scores	Z-Scores	T-Scores
1	-2.66	23	16	-.52	45
2	-2.52	25	17	-.37	46
3	-2.38	26	18	-.23	48
4	-2.24	28	19	-.09	49
5	-2.09	29	20	.05	50
6	-1.95	30	21	.19	52
7	-1.81	32	22	.33	53
8	-1.66	33	23	.48	55
9	-1.52	35	24	.62	56
10	-1.38	36	25	.76	58
11	-1.23	38	26	.91	59
12	-1.09	39	27	1.05	60
13	-.95	40	28	1.20	62
14	-.81	42	29	1.33	63
15	-.66	43	30	1.48	65

M=19.64; SD=6.96

Table 43

Z-Scores and T-Scores Equivalents of the raw scores for SVITU (N = 1080)

Raw Scores	Z-Scores	T-Scores
1.	-4.472	5
2.	-4.416	6
3.	-4.359	6
4.	-4.303	7
5.	-4.246	7
6.	-4.189	8
7.	-4.133	8
8.	-4.076	9
9.	-4.020	9
10.	-3.963	10
11.	-3.906	11
12.	-3.850	11
13.	-3.793	12
14.	-3.737	12
15.	-3.680	13
16.	-3.623	14
17.	-3.567	14
18.	-3.510	15
19.	-3.453	16
20.	-3.397	16
21	-3.340	17
22.	-3.284	17
23.	-3.227	18
24.	-3.170	18
25.	-3.114	19

Cont...

Table 43

Z-Scores and T-Scores Equivalents of the raw scores for IIT (N = 1080)

Raw Scores	Z-Scores	T-Scores
1.	-4.472	5
2.	-4.416	6
3.	-4.359	6
4.	-4.303	7
5.	-4.246	7
6.	-4.189	8
7.	-4.133	8
8.	-4.076	9
9.	-4.020	9
10.	-3.963	10
11.	-3.906	11
12.	-3.850	11
13.	-3.793	12
14.	-3.737	12
15.	-3.680	13
16.	-3.623	14
17.	-3.567	14
18.	-3.510	15
19.	-3.453	16
20.	-3.397	16
21	-3.340	17
22.	-3.284	17
23.	-3.227	18
24.	-3.170	18
25.	-3.114	19

Cont...

Score	Z-Score	T-Score
26.	-3.057	19
27.	-3.001	20
28.	-2.944	20
29.	-2.887	21
30.	-2.831	21
31.	-2.77	42
32.	-2.718	23
33.	-2.661	23
34.	-2.604	24
35.	-2.548	24
36.	-2.491	25
37.	-2.435	25
38.	-2.378	26
39.	-2.321	27
40.	-2.265	27
41.	-2.208	28
42.	-2.151	28
43.	-2.095	29
44.	-2.038	29
45.	-1.982	30
46.	-1.925	31
47.	-1.868	31
48.	-1.812	32
49.	-1.755	32
50.	-1.699	33
51.	-1.642	33
52.	-1.585	34
53.	-1.529	35
54.	-1.472	35
55.	-1.416	36

Cont...

Score	Z-Score	T-Score
56.	-1.359	36
57.	-1.302	37
58.	-1.246	37
59.	-1.189	38
60.	-1.132	38
61.	-1.076	39
62.	-1.019	40
63.	-.963	40
64.	-.906	41
65.	-.849	41
66.	-.793	42
67.	-.736	42
68.	-.680	43
69.	-.623	43
70.	-.566	44
71.	-.510	45
72.	-.453	45
73.	-.397	46
74.	-.340	46
75.	-.283	47
76.	-.227	47
77.	-.170	48
78.	-.114	49
79.	-.057	49
80.	-.006	50
81.	.055	50
82.	.112	51
83.	.169	51

Score	Z-Score	T-Score
84.	.225	52
85.	.282	53
86.	.338	53
87.	.395	54
88.	.452	54
89.	.508	55
90.	.565	55
91	.622	56
92.	.678	56
93.	.735	57
94.	.791	58
95.	.848	58
96.	.905	59
97.	.961	59
98.	1.018	60
99.	1.074	60
100.	1.131	61
101.	1.188	62
102.	1.244	62
103.	1.301	63
104.	1.357	63
105.	1.414	64
106.	1.471	64
107.	1.527	65
108.	1.584	66
109.	1.640	66
110.	1.697	67
111.	1.754	67

Raw Scores	Z-Scores	T-Scores
112.	1.810	68
113.	1.867	68
114.	1.924	69
115.	1.980	70
116.	2.037	70
117.	2.093	71
118.	2.150	71
119.	2.207	72
120.	2.263	72
121.	2.320	73
122.	2.376	73
123.	2.433	74
124.	2.490	75
125.	2.546	75
126.	2.603	76
127.	2.659	76
128.	2.716	77

M=80.01; SD=17.06

DISCUSSION

CHAPTER-4

DISCUSSION

An Sajjad Verbal Intelligence Test in Urdu for adolescents (SVITU) has been developed, validated and standardized with a purpose to measure and assess the general Intelligence of the adolescents and adults within the age (17-20 years) of Pakistan in the national language, (Urdu). Effort has been made to make SVITU and its sub-tests psychometrically valid and reliable.

The SVITU consists of four sub-tests: Vocabulary test, Verbal Reasoning test, Numerical Reasoning test, and Information test. The total number of items of the test is 128 and time limit is 40 minutes.

The item analysis (tables 4-5) was carried out to select the best items. The items having discrimination power within the range (.50 to .80) were considered highly effective as they seem to discriminate between high and low achievers on the test. Items showing discrimination power less than .30 were discarded from the item pool. Results presented in table 4 show that 128 items fall within the desired range of discrimination (.3 and above). As far as the difficulty levels of the items is concerned, the items falling within the range (.41-.70) were retained, since they are neither very easy nor very difficult.

Tables 4-5 reveal that items in each sub-test give a reasonable distribution of ranges of discrimination power and difficulty level. Most of the items possess sound psychometric bases, except the items of Verbal Reasoning Test, which do not seem as effective as the items of the other sub-tests. Verbal Reasoning Test was found to be the easiest out of the four sub-tests, while other sub-tests were found reasonably effective measure of intelligence.

Item discrimination analysis of Vocabulary Test (table 14) indicates that high-scoring group scored significantly high on 40 items, whereas the remaining two items (7 and 12) do not significantly differentiate between the performances of high and low scoring groups. This finding demonstrates that out of 42 items, 40 items are quite effective in discriminating between high and low scorers.

Item discrimination analysis of Verbal Reasoning Test (table 15) reveals that high-scoring group scored significantly high on 17 items. Whereas the remaining three items (2,3, and 7) do not significantly differentiate between the performances of high and low scoring groups. This finding demonstrates that 17 out of 20 items are quite effective in discriminating between high and low scorers.

Item discrimination analysis of Numerical Reasoning Test (table 16) shows that high-scoring group scored significantly high on 35 items. Whereas the remaining one item (4) does not significantly differentiate between the performances of high and low scoring groups. This finding demonstrates that out of 36 items, 35 are quite effective in discriminating between high and low scorers.

Item discrimination analysis of Information Test (table 17) indicates that high-scoring group scored significantly high on all the 30 items. This finding shows that all items of this sub-test are quite effective in discriminating between high and low scorers. This finding about Information Test may be due to the fact that the population sample over which the test was administered was appearing before the entry test in Army. They might have, already, prepared their general knowledge for that purpose, which resulted in the good performance in this sub-test.

Item discrimination analyses of the items of four sub-tests of SVITU establish the validity of each item in measuring intelligence, besides ascertaining their discriminatory power.

Tables 18-19 shows that reliability coefficients of three applied methods for SVITU and its sub-tests are highly significant. These results suggest a higher level of reliability both in terms of internal consistency and

temporal stability of results. The reliability results show that the test is a reliable instrument for measuring general intelligence.

The validity coefficients of the SVITU and its sub-tests (tables 20-21) indicate significant correlation between SVITU and Army Intelligence Test (AIT) and between SVITU and college marks. These findings establish the construct and concurrent validity of the test and demonstrate the validity of SVITU as a valid measure of general intelligence.

Results presented in table 20 show that the coefficient of correlation between AIT with the Vocabulary Test , Verbal Reasoning Test, Numerical Reasoning Test, Information Test and SVITU are positive and significant ($p<.001$). According to Anastasi (1990), these correlations, if moderately high can be cited as an evidence that the new intelligence test measures approximately the same general area of behavior as other tests designed by the same name as *Intelligence Tests*. These findings clearly establish the construct validity of SVITU.

An inspection of the Table 20 shows that SVITU and its sub-tests demonstrate low correlation with the nonverbal part of Army Intelligence Test as compared to its verbal part. The results show that verbal test of intelligence correlate high with the tests measuring the verbal ability as compared to the nonverbal intelligence tests. The results support the theoretical assumption of Campbell and Fiske(1959), who in their Multitrait- Multimethod Approach

noted that if a number of methods are used to measure more than one trait or construct, the correlations among them will take the form of a multitrait-multimethod matrix which is useful in assessing construct validity. According to them the convergence of different methods to measure the same trait raises the confidence of the test developer in the construct and thus establishes the construct validity.

Table 21 shows that the correlation between college marks and scores of IIT and its sub-tests are statistically significant ($p<.001$). These results support the study by Crano, Denny and Campbell (1972), who inferred that correlation of intelligence tests taken at one point of time and achievement measure taken at a later point of time tend to be higher than the two measures taken in reverse order of times. These results are in accordance with the theoretical assumption of Murphy and Davidshofer (1998), that is, theoretically a correlation could range in absolute value from 0.0 to 1.0, whereas in practice, most validity coefficients tend to be fairly small. A good, carefully chosen test is not likely to show a correlation greater than .5 with an important criterion and, infact, validity coefficients greater than .3 are not all common in applied settings. Schmidt, Hunter, & Pearlman (1981) in one of the studies found validity coefficients of .19 .24 and.21 for job grades with verbal ability, quantitative ability and reasoning ability respectively. Guion (1991) noted that validity of measurement is not always necessary to guarantee validity for decision. Although both sets of scores seem to be going

in the same direction, it should not be inferred that any one who does good on a traditional college examination, does equally good on intelligence test. However, one can conclude that educational achievement has nothing much to do with the understanding of the text knowledge; rather one can say that school achievement is not dependent upon only intelligence. Modern researches on general intelligence have proved validity generalization aimed at forecasting educational outcomes, occupational training, and work performance. For some benchmarks, general cognitive ability covaries .70 to .80 with academic achievements, and .40 to .70 with military training assignments, if intelligence test is administered before the two mentioned events.(Brody, 1992; Gottfredson, 1997; and Jensen,1998.).

The coefficient of correlation between SVITU and its sub-tests (table22) are also highly significant ($p<.001$). The results demonstrate the internal consistency of SVITU and its sub-tests and indicate that all the sub-tests measure some universal or general factor and similar mental functions. These positive correlation indicates that all the verbal items included in the sub-tests are loaded on general intelligence factor g and crystallized g_f. The correlations of four sub-tests with the total score also provide further evidence of the consistency of the test across individual components.

The results regarding grade differentiation (tables 23-28) reveals that only Verbal Reasoning Test showed no significant difference between the

grades(9^{th} to 13^{th}), whereas other subtests and SVITU demonstrated significant degree of difference between the grades. The reason of not receiving significant difference on Verbal Reasoning Test may be due to the fact that Verbal Reasoning Test has been proved to be the easiest test of all the subtests.Tables23-27 showing the means and standard deviations of SVITU and sub-tests demonstrate a general decline in the scores with decline in the age. This finding confirms Sternberg (1997) study in which comparison between the intelligence of young and the old adults indicates that average performance of young adults tend to be higher.

In view of the statistical evidence presented in the above paragraphs, it can be concluded that items included in SVITU have demonstrated well their validity and when presented in the form of a test can prove as an effective measure of intelligence, as it can differentiate adequately between individuals varying in intellectual abilities.

Generally there exists no significant gender difference in general intelligence, however, differences are observed when the factors making up the intelligence are broken into its different parts. The data of present research (table 30) show that females did better than males on Vocabulary test of SVITU. These findings confirm earlier researches (e.g Weschler ,1938) who believes that females tend to be superior than males in rote memory, vocabulary and verbal fluency.

In another study, Feingold (1992) reported that males varied than females in quantitative reasoning, spatial visualization, spelling and general knowledge. Because these sex differences in variability were coupled with corresponding sex difference in means.

The difference between males and females on the Information Test (table 29) may be attributed to the gender specific socialization practices of our society. Boys performed better on Information Test which may be attributed to the fact that they generally are more exposed to the general happenings of common life and are more acquainted with current affairs. Another reason for the good performance of boys on information test seems to the fact that the sample comprised of boys appearing before Army Selection and Recruitment Centers for selection as cadet in Army. They were all well aware that they have to be evaluated in terms of their knowledge about current affairs and general information.

Table 29 reveals that the items of Vocaulary test and Numerical Reasoning Test are likely to favor the females as compared to males while Information test may favour the experiences of males than females.

The differences in the intelligence scores of the students of English and Urdu medium schools as shown in table 30, may be attributed to their

differences in social and educational background. Students belonging to Urdu medium schools mostly belong to the lower socioeconomic group with poor educational facilities. Besides, Urdu medium schools provide with much lesser physical, social and educational facilities than English medium schools, which are usually better equipped with all sort of facilities.

Table 30 reveals that students of Urdu medium schools did better on Vocabulary test and on Information test. This difference of performance may be attributed to the fact that students are in advantageous position as compared to English medium students as far as command of Urdu language is concerned. However, English medium Students did slightly better on Verbal Reasoning Test and Numerical Reasoning Test. This may be due to the fact that better educational facilities might have raised their reasoning power as compared to Urdu medium students.

As far as difference in the test scores of students having urban and rural backgrounds is concerned, the results (table 31) shows significant difference ($p<.01$)on Vocabulary test and Verbal Reasoning test. Here rural persons did better due to the more acquaintance with the language of Urdu. This may be attributed to the fact that educational facilities in Urdu language might have raised their reasoning power as compared to the students with urban background.

Significant differences were found in the performances of science and arts students(table 32). Science students did better in all the four sub-tests and total test as compared to arts students. This difference may be attributed to educational facilities available to science students. Moreover, the science students are more analytical and supposed to work hard to get through their examination and competition for admission to professional colleges after completing their intermediate level of education, hence their scores are likely to be better than arts students.

Regarding the performances of the sons of government servants and sons of businessmen, results (table 33) show that the former group has a slight edge over the latter on Numerical Reasoning test and Information test. This difference may be attributed to the fact that the sons of government servants are more serious natured, task oriented and have learned to rely on their own capabilities. No significant difference between the two groups was found on other sub-tests.

To interpret the percentile norms, after having a look at tables 34-38, it is useful to describe a person as superior if his score lies at or above the 95^{th} percentile for the normative group, above average, when his score lies at or above 75^{th} percentile, average when his score lies between 25th-75th percentiles, below average, when his score lies at or above 10^{th} percentiles, and defective, when his score lies at or above 5^{th} percentiles.

Tables 39-43 demonstrates that the obtained T-scores of SVITU and sub-tests fall within the acceptable range of 20-80 (Murphy & Davidshoffer, 1998).

Limitations

There appeared a lot of difficulties and practical problems in developing an indigenous intelligence test due to dearth of previous research in this field and non-availability of a working model. Despite the fact that research studies have been conducted on development, adaptation and validation of personality, intelligence, ability, aptitude, and achievement tests and related issues, there is little work found on the development of indigenous intelligence tests, especially for adolescents and adults. Most of the psychometric research in Pakistan had been in English language.

In view of the previous researches done in the field of intelligence testing, the study of development of Indigenous Intelligence Test in Urdu Language, can be considered as a first effort based on the scientific principles of test development.

Since this study was carried out mainly on male students, who reported to Army Selection and Recruitment Centers for the induction to Army, therefore the newly developed test, SVITU is recommended to be used for the same type of population.

Future Recommendations

1. To ascertain the psychometric characteristics of items, some other statistical calculations like point biserial correlation for item discrimination can be carried out.
2. To provide additional support to the validity of the test, the sample size of the study can be enlarged by including students from both genders, rural areas and other urban settings belonging to different social, cultural and educational backgrounds.
3. In order to increase the reliability of the test, items found not very strong regarding item discrimination and item difficulty indices can be improved and some new items, especially in Verbal Reasoning Test should be added. Moreover, some other sub-tests of verbal ability can also be added in SVITU.
4. Researches can be carried out to validate the test for general population including boys and girls. Similarly studies can be conducted to develop norms for all the four provinces of Pakistan.
5. Cross-cultural comparisons can be carried out between the countries of South-Asia, like Bangladesh and India, where Urdu language is spoken, read, and understood.

SUMMARY

The main purpose of this study was to develop, validate and standardize a Group Verbal Intelligence Test in Urdu for adolescents to measure and assess the general intelligence of the students who had completed grade 12 and are within the age range 17 to 20 years.

To generate the items, the available material regarding verbal intelligence tests was reviewed. The selected test material measures different aspects of general intelligence like vocabulary, verbal reasoning, numerical reasoning, information, etc. The items were developed according to the stipulated areas of content. The ideas for the items were also borrowed from standardized verbal intelligence tests, like Otis-Lennon Test. Initially a pool of 250 items falling into four sub-tests was developed. These sub-tests were: Vocabulary test, Verbal Reasoning test, Numerical Reasoning test, and Information test. The basic reason for including four sub-tests was to make the test as varied as possible so as to measure different aspects of one's intellectual functioning. Initial item pool was tried out on a sample of 60 students of grades 11 and 12 of PAF College Rawalpindi, within the age range of 17 to 20 years.

After pre-testing, 230 items were retained. The four sub-tests were then administered on a sample of 200 candidates of PMA Long Course. The

items were analyzed with reference to internal consistency, difficulty level and discrimination power. Keeping in view the three indices, 153 items were retained. These items were rearranged according to their difficulty level and were administered to another sample of 200 candidates of PMA Long Course for second item analysis. Thus, the items were reduced to 128, which were rearranged according to their descending order of difficulty level.

Before administering the sub-tests for main study, average time required for SVITU and its sub-tests was calculated in a separate study. The final test was then administered to a sample of 535 candidates of PMA Long Courses. The sample was taken from the Army Selection and Recruitment Centers of Pakistan. The purpose of the main study was to determine different psychometric characteristics of the test.

The reliability of the SVITU was determined by Kuder Richardson method, Split-half method, and test-retest method. The results demonstrate high reliability of the test both in terms of internal consistency as well as temporal stability of the results.

The validity of the test was determined by four procedures. The construct validity of the test was studied by correlating it with Army Intelligence Test. The concurrent validity was determined by correlating it with the college marks of the students in their annual examination. Both the

items were analyzed with reference to internal consistency, difficulty level and discrimination power. Keeping in view the three indices, 153 items were retained. These items were rearranged according to their difficulty level and were administered to another sample of 200 candidates of PMA Long Course for second item analysis. Thus, the items were reduced to 128, which were rearranged according to their descending order of difficulty level.

Before administering the sub-tests for main study, average time required for IIT and its sub-tests was calculated in a separate study. The final test was then administered to a sample of 535 candidates of PMA Long Courses. The sample was taken from the Army Selection and Recruitment Centers of Pakistan. The purpose of the main study was to determine different psychometric characteristics of the test.

The reliability of the IIT was determined by Kuder Richardson method, Split-half method, and test-retest method. The results demonstrate high reliability of the test both in terms of internal consistency as well as temporal stability of the results.

The validity of the test was determined by four procedures. The construct validity of the test was studied by correlating it with Army Intelligence Test. The concurrent validity was determined by correlating it with the college marks of the students in their annual examination. Both the

indices yield evidence of high validity of the test. Inter correlations of sub-tests and their correlations with the total test revealed internal consistency of the sub-tests and established the construct validity of the test.. One way ANOVA was computed to study if SVITU and its sub-tests demonstrate the grade differentiation (grades 9^{th} to 13^{th}). The results confirmed the validity of the test as a measure of general intelligence.

To study the significance of differences between various groups, t-test was applied. The results show significant differences between the performance of male and female students, between students of English/Urdu medium schools, between urban and rural students, between science and arts students, and between the sons of government servants and businessmen. Percentile norms and T-Scores were developed for SVITU and its sub-tests.

REFERENCES

REFERENCES

Addair, J. G., & Puhan, B. N. (1993). Indigenization of psychology: Empirical assessment of progress in India research. *International journal of Psychology, 28*, 149-169.

Adkin, D.C. (1974). *Test construction: development and interpretation of achievement tests. (2nd Ed).* Ohio: A Bell & Howell Company.

Ahmad, I. (1986). Initial development and validation Academic Self Concept Scale. *Pakistan Journal of Psychological Research, 1 (3-4),* 43-50.

Ahmad, I. (1987). Initial psychometric evaluation of Urdu version of California Psychological Inventory (CPI).*Pakistan Journal of Psychological Research, 1* (1-2), 3-16.

Aiken, L. A. (1988) *Psychological testing and assessment.* London: Allyn and Bacon, Inc.

Ain, Q. (1985). A validation study of Cattlell's Culture Fair Intelligence Test on Pakistani children. New Directions in Pakistani Psychology: *Proceedings of the 5th Conference of Pakistan Psychological Associations.* Z. A. Ansari; P. N. Tariq; R. M. Zaman; S. Chowdhri, and S. Watanabe, (Editors). Islamabad: 155-161.

Anastasi, A. (1970).The formation of psychological traits. *American Psychologist, 25,* 899-910.

Anastasi, A. (1976). *Psychological Testing (4th ed.).* New York: Collier Macmillan International Edition.

Anastasi, A. (1990). *Psychological Testing (6th ed.).* New York: Macmillan Publishing Company.

Anderson, J.R. (1982). Acquisition of cognitive skill. *Psychological Review, 89,* 369-406.

Ansari, Z. A. (1976). Creativity, intelligence and academic achievement. *Pakistan Journal of Psychology , 9* (3-4) 17-26. -44.

Ansari, Z. A. (1977). University examinations in terms of the Taxonomy of Educational Objectives. *Pakistan Psychological Studies, 2*, 29-42 .

Ansari, Z. A. (1979). *Problems of psychological research in Pakistan.* Sectional presidential address at fourth annual session. Pakistan Psychological Association, Karachi.

Ansari, Z. A. (1984). Achievement oriented attitude and academic performance of Pakistani high school students. *Pakistan Psychological Studies, Vol 5,*21-28.

Ansari, Z. A., & Chowdhri, S. (1988). Development of a questionnaire for assessing the study problems of Pakistani high school students. *Pakistan Journal of Psychological Research, 3 (3-4)*, 1-9.

Ansari, Z. A., & Iftikhar, M. (1988).Validity of Raven's Standard Progressive Matrices for urban and rural school children in Pakistan, Part I: Basic facts. *Psychology Quarterly, 19*, 14-27.

Ansari, Z. A., Tariq, P. N., & Iftikhar, M. (1990). *Educational Ability Test Level 5: Development and validation.* Islamabad: National Institute of Psychology, 48p.

Azuma, I. (1984). Psychology in a non-Western country. *International Journal of Psychology, 19*, 45-56.

Baker, K. H. (1939). Item validity by the analysis of variance: An outline of method. *Psychological Record, 3*,242-248.

Baltes, P. B., Dittmann-Kohli, F., & Dixon, R. A. (1984). *New perspectives on the development of intelligence in adulthood: Toward a dual-process conception and a model of selective optimization with compensation.*. In P.B Baltes & O. G. Brim, Jr (Eds.), *Life-span development and behavior, 6* New York: Academic Press.

Beilin, H. (1923). *The new functionalism and Piaget's program.* In E.K Scholnick (Ed,), *New trends on conceptual representation.* Hillsdale, NJ: Erlbaum.

Bennett, G. K., Seashore, H. G., & Wesman, A. G. (1947) *Differential Aptitude Test.* New York: Psychological Corporation.

Berry, J. W., & Dasen, P. R. (1974). *Cultural and Cognition: Readings in Cross-Cultural Psychology*. London: Methuen.

Binet, A. (1890). Perceptions d'enfants. *Revue Philosophique, 30,* 582-611.

Binet, A., & Simon, T. (1905). Mehodes nouvelles pourle diagnostic du niveau intellectuel des anormaux. *L'Annee Psychologique, 11*, 191-244

Binet, A., & Simon, T. L(1909).'intelligence des imbecilies. *L'Annee Psychologique, 15,* 1-147

Birnbaum, A. (1968). *Some latent trait models and their use in inferring an examinee's ability.* In F. M. Lord & M. R. Novick, *Statistical theories of mental test scores* (Pt. 5). Reading, Mass: Addison-Wesley.

Bock, R. D., & Wood, R. (1971). Test theory. *Annually Review of Psychology, 22,* 193-224.

Boring, E. G. (1923a). Intelligence as the tests test it. *New Republic, 35,* 35-37. paedo-psychiatric findings and EEG variables in well functioning children of ages 5-16 years. *Scandinavian Journal of Psychology, 18,* 140-147.

Boring, E.G. (1923b). *A history of experimental psychology. (Revised ed.).*New York: Appleton-Century Crafts.

Boring, E. G. (1950). *A history of experimental psychology (rev. ed.)* New York: Appleton-Century-Crafts.

Born, M., & Flier, D. V (1987). Cross-cultural comparison of sex related differences in intelligence tests. A meta analysis. *Journal of Cross-cultural Psychology. 18. (3)..*

Brigham, C. C. (1930). Intelligence test of immigrant groups. *Psychological Review, 37,* 158-165.

Brody, N. (1992). *Intelligence. (2nd ed).* SanDiego C A: Academic Press.

Brown, W., & Thomson, G. (1921). *The essentials of mental measurement.* Cambridge: Cambridge University Press.

Bruno, F. J. (1986). *Dictionary of key words in psychology*. 114-115.Routledge & Kegan Paul London: Boston and Henley.

Burt, C. (1938). Factor analysis by sub-matrices. *Journal of Psychology, 6,* 339-375.

Butcher, H. J. (1973). *Intelligence and creativity. In Kline, P. (Ed.) New Approaches in Psychological Measurement*. London: Wiley.

Campbell, D. P. (1965). A cross-sectional and longitudinal study of scholastic abilities over twenty-five years. *Journal of Counseling Psychology, 12*, 55-61.

Carroll, J. B. (1945). The effect of difficulty and chance success on correlations between items or between test. *Psychometrika, 10*, 1-19.

Carroll, J. B. (1980). Individual difference relations in psychometric and experimental cognitive tasks. (Report. No. 163), University of North Carolina.

Carroll, J. B. (1983). *Individual differences in cognitive abilities.* In Irvine, S. H., & Berry, J. W. (Eds). Pp. 213-235.

Carroll, J. B. (1993).*Human cognitive abilities:a summary of factor analytic studies.* Cambridge: Cambridge University Press.

Cattell J. M. (1890). Mental tests and measurements. *Mind, 15,* 373,-381.

Cattell J. M., & Farrand, L (1896). Physical and mental measurement of the students of Columbia University. *Psychological Review, 3*, 618-648.

Cattell R. B. (1965). *Higher order factor structures and reticular vs. hierarchical formulae for their interpretation.* In C. Banks & P. L. Broadhurst (Eds.) London: University of London Press.

Cattell, J B., & Maxwell, S. E. (1979). Individual differences in cognitive abilities. *Annual Review of Psychology, 30,* 603-640.

Cattell, J. M. (1886a). Psychometrische Untersuchungen. *Philosophische Stdien, 3,* 452-492.

Cattell, J. M. (1886b). The time taken up by cerebral operations. Mind, 11, 222, 242, 377-392, 524-538.

Cattell, J. M. (1885). Uber die Zeit der Erkennung und Benennung von Schriftzeichen, Bildern und Farben. *Philosophische Studien* 2,635-650. Translated in A. T. Poffenberger (Ed.), James Mckeen Cattell: Man of science. Vol I: Psychological research. Lancaster, Pa Science Press, 1947.

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin, 40,* 153-193.

Cattell, R.B. (1963). Theory of fluid and crystallized intelligence; A critical experiment, *Journal of Education Psychology,54,* 1-22.

Cattell, R.B. (1971). *Abilities: Their structure, growth, and action*, Boston: Houghton Mifflin .

Cattell, R.B., & Butcher, J. (1968). *The prediction of achievement and creativity.* Indianapolish.Ind:Bob Merrill, .

Charlesworth, W. R. (1983). *An ethological approach to cognitive development. In C.J. Brainerd (Ed). Recent advances in cognitive-developmental theory: Progress in cognitive development research.* New York: Springer.

Chi, M.T.H., Glaser, R., & Rees, E. (1983) *Expertise in problem solving.* In R.J Sternberg (Ed.) *Advance in the psychology of human intelligence (vol-1)* Hillsdale, NJ: Erlbaum.

Cook, T.D., & Campbell, D.T. (1979). *Quasi - experimentation.* Chicago: Rand McNally.

Cornelius. S. W., & Kenney, S. R. (1982) *Academic and everyday intelligence in adulthood and old age*, Unpublished manuscript, Department of Human Development and Family Studies, Cornell, New York: University, Ithaca.

Crano, W.D. (1974). *Causal analyses of the effects of socioeconomic status and initial intellectual endowment of patterns of cognitive development and academic achievement.* In D.R, Green (Ed), The aptitude-achievement distinction. Monterey, California: CTB/McGraw - Hill.

Crano, W.D., Denny, D.A., & Campbell, D.T. (1972). Does intelligence cause achievement? A cross-lagged panel analysis, *Journal of Educational Psychology, 63,* 258-275.

Cremin, L. A. (1981). *The transformation of the school: Progressivism in American education, 1876-1957.* New York: Knopf.

Cronbach, L. J., Rajaratnam, N., & Gleser, G. C. (1963). Theory of genralizability: A liberation of reliability theory. *British Journal of Mathematical and Statistical Psychology, 16,* 137-163.

Cronbach, L. J., Rajaratnam, N., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles.* New York: Wileym.

Cronbach, L. A. (1961). *The transformation of the school: Progressivism in American education, 1876-1957.* New York: Knopf.

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

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of genralizability for scores and profiles.* New York: Wiley.

Dar, I. S. (1982, June). Test development for vocational counseling in Pakistan: Problems and issues. *Pakistan Journal of Psychology, 3* (1-2), 25-31.

Demming, J.A., & Pressey, S. L. (1957). Tests indigenous to the adult and older year, *Journal of Counseling Psychology, 4*, 144-148.

Deregowski, J. B. (1980). *Illusions, patterns, and pictures: A cross-cultural perspective.* London: Academic Press.

Dorner D. (1982). *The ecological conditions of thinking.* In D.R Griffin (Ed), *Animal mind Human mind* (Dahlem Konferenzen). New York: Springer.

Dorner, D (1983). *Heuristic and cognition in complex systems. In R. Groner. M Groner, & W.F. Bischof (Eds), Methods of heuristics.* Hillsdale. Hillsdate, NJ: Erlbaum,

DuBois, P. H. (1970). *A history of psychological testing.* Boston: Allyn & Bacon.

Ebbinghaus, H. (1897). Uber eine neue Methode zur Prufung geistiger Fahigkeiten und ihre Anwendung bei Schulkindern. *Zeitschrift fur angewandte Psychologie, 13*,401-459.

Edelstein, W., & Noam, G. (1982). Regulatory structures of the self and " Postformal" stages in adulthood. *Human Development, 6*, 407-422.

Eysenck, H. J. (1967). Intelligence assessment. A theoretical and experimental approach. *British Journal of Educational Psychology, 37,* 81-98.

Eysenck, H. J. (1979). *The structure and measurement of intelligence.* NewYork:Springer-Verleg.

Eysenck, M. W. (1984). *A handbook of cognitive psychology.* London: Erlbaum.

Fechner, G. (1860). Elemente der Psychophysik, Leipzig: Breitkopf & Hartel.

Ferguson, G. A. (1942). Item selection by the constant process. *Psychometrika, 7,* 19-29.

Ferguson, G.A. (1954a). On learning and human ability, *Canadian Journal of Psychology, 8,* 95-112.

Ferguson, G.A. (1954b). On transfer of the abilities of man. *Canadian Journal of Psychology, 10*,121-131.

Finney, D. J. (1944). The application of probit analysis to the results of mental tests. *Psychometrika, 9*, 31-39.

Flanagan, J. C. (1948). The aviation psychology program in the Army Air Forces (Rep. I, Army Air Forces Aviation Psychology

Flanagan, J. C. (1960). *Flanagaun aptitude classification test*. Chicago: Science Research Associates, 1951-1960.

Flanagan, J. C., & Cooley, W. W. (1966). *Project talent: One year follow-up studies*. Pittsburgh: University of Pittasburgh, Project Talent.

Flanagan, J. C., Dailey, J. T., Shaycoft, M. F., Gorham, W. A., Orr D. B., & Goldberg, I. (1962). *Design for a study of American youth, I: The taklents of American youth.* Boston: Houghton-Mifflin.

Fogiel, M. (1980). *Psychological Problem Solver (rev.ed).* New Jersy: Research and Education Association.

Freeman, F. S. (1965). *Theory and Practice of Psychological Testing (3rd.ed.).* New Delhi: Oxford & IBH Publishing Company.Private Limited.

Galton, F. (1869). *Hereditary genius: An inquiry into its laws and consequences.* New York: Appleton.

Galton, F. (1870). *Inquiries into human faculty and its development.* New York: Appleton.

Gardezi, A. H. (1994). *Development of a non-verbal intelligence test for adolescents.* Unpublished M. Phil. dissertation. National Institute of Psychology, Islamabad.

Gardner, H. (1993). *Multiple intelligence: The theory in practice.* NewYork: Basic Books.

Garrett, H.E. (1946).A developmental theory of intelligence, *American Psychologist, 1*, 372-378.

Goddard, H. H. (1908). The Binet and Simon tests of intellectual capacity. *The Training School 5*, 3-9.

Gottfredson, L. S. (1997). Intelligence and social policy [special issue]*Intelligence,24,* 2-8.

Gregory, R. J. (1996). *Psychological Testing: History, principles and applications. (2nd Ed).* London: Allyn &Becon Inc.

Groundland , N.E. (1990) *Measurement and Evaluation in teaching.* London: Macmillen Publishing Company.

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

Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.

Guilford, J. P. (1971). *The nature of human intelligence*. New York: McGraw-Hill.

Guilford, J. P. (1988). Some changes in the structure of intellect model. *Educational &Psychological measurement,48,*1-4.

Guilford, J. P., & Hoepfner, R. (1971). *The analysis of intelligence*. New York: McGraw-Hill.

Guilford, J. P., & Zimmerman, W. S. (1956). *Guiford-Zimmerman Aptitude Survey.* Orange, California: Sheridan Psychological Services, 1947-1956.

Gulliksen. H. (1950). *Theory of mental test.* New York: Wiley.

Guttman, L. (1941). *The quantification of a class of attributes: A theory and method for scale construction.* NewYork: Social Sciences Research Council.

Guttman, L. (1965). The structure of relations among intelligence tests, *Proceeding of the 1964 invitational conference on testing problems.*_Princeton, N.J: Educational Testing Service.

Guttman,L. (1969). Integration of test design and analysis. *Proceedings of the 1969 invitational conference on testing problems,* Princeton, N.J: Educational Testing Service.

Harnquist, K. (1968). Relative changes in intelligence from 13 to 18. I.Background and methodology. II. Results. *Scandinavian Journal of Psychology, 9,* 50-82.

Hasan, R. (1981). A comparative study of the Pakistani bilingual and monoglot school children's performance in verbal and non-verbal test. *Pakistan Journal of Psychology, 12 (3-4)*, 3-12.

Hassan, P. (1986). *Test developed and standardized for children at the board of secondary education Karachi.* Paper presented at the seminar on child development. National Institute of Psychology, Islamabad.

Henmon, V. A. C., & Nelson, M. J. (1935). *Henmon-Nelson test of mental ability.* Boston: Houghton-Mifflin 1932-1935.

Hofstadter, R. (1944). *Social Darwinism in American thought.* Philadelphia: University of Pennsylvania Press.

Horn, J.L (1970). *Organization of data on life-span development of human abilities.* In L.R. Goulet & P.B Baltes (Eds), *Life -span development psychology; Research and theory.* New York: Academic Press.

Horn, J.L (1982), *The theory of fluid and crystallized intelligence in relation to concepts of cognitive psychology and aging in adulthood.* In F.I.M. Craik & S, Trehub (Eds), Aging and cognitive process New York: plenum Press.

Horn, J.L. (1976). Human abilities: A review of research and theory in the early 1970's. *Annual Review of psychology, 27,* 437-485.

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology, 24,* 417-441,

Hoyer, W.J. (1984). *Aging and the development of expert cognition.* In T.M Schlechter & M.P. Toglia (Eds). *New directions in cognitive science.* Norwood, New Jersy: Ablex.

Hoyty, C. (1941). Test reliability estimated by analysis of variance. *Psychometrika, 6,* 153-160.

Humphreys, L.G., & Taber. T,(1970). Postdiction study of the graduate record examination and eight semesters of college grades.*Journal of Educational Measurement, 10,*179-184.

Hunt, E. B. (1976). varieties of cognitive power. In Resnik, R. B. (Ed.).

Hunt, E. B. (1978). Mechanics of verbal ability. *Psychological Review, 85,* 109-130.

Hunt, E, (1983) One the nature of intelligence. *Science, 219,*141-146.

Husen, T. (1951).The influence of schooling upon IQ. *Theoria,17,*61-88.

Hussain, S. S. (1992). *Development of a group verbal intelligence test in Urdu for high school students.* Unpublished M. Phil. Dissertation, Islamabad: National Institute of Psychology, 66p.

Imam, A., & Munaf, S. (1988). Birth order and intelligence. *Pakistan Journal of Psychology, 19 (1-4),* 27-35.

Ismail. Z, & Mehmood, L. (1986). Social class and sex differences in Pakistan on the Raven's Standard Progressive Matrices. *Pakistan Journal of Psychology, 17 (3-4),* 23-35.

Israr, M. (1979). *Development of the concept of number in primary school children.* Islamabad: National Institute of Psychology, 8p.

Israr, M. (1980). *Science and mathematics curriculum in Pakistani primary school: A cognitive demand analysis.* Islamabad: National Institute of Psychology, 81p.

Israr, M. (1982a). *Patterns of language used in parent-child interaction in Pakistan society.* Islamabad: National Institute of Psychology, 28p.

Israr, N (1982b). *A Cognitive Development Task Battery for primary school children of Pakistan.* Islamabad: National Institute of Psychology. 51p.

Israr, M. (1985). Cognitive development of Pakistani primary school children: Salient findings of a research project; New Directions in Pakistani psychology: *Proceedings of the 5th Conference of Pakistan Psychological Associations. Ansari, Z. A.; Tariq, P.N.; Zaman, R.M.; Chowdhri, S. and Watanabe, S. (Editors).* Islamabad: PPA, 67-75.

Israr, M. (1988). *A psychological interpretation of mathematics learning problems among secondary schools students.* Unpublished M. Phil. Dissertation, Islamabad: National Institute of Psychology, 78 + 15p.

Israr, N., & Pervez, M. (1986). Cognitive development task battery for class 1 primary school children. *Pakistan Journal of Psychological Research, 1 (3-4),* 1-15.

Jensen, A.R. (1970). *Hierarchical theories of mental ability. In W.B. Dockrell (Ed), On intelligence.* London: Methuen.

Jensen, A. R. (1980). *Bias in mental testing.* New York: Free Press.

Jensen, A. R. (1998). *The g factor.* NewYork: Praeger.

Joreskog, K. G., & Lawley, D. N. (1968). New methods in maximum likelihood factor analysis. *British Journal of Mathematical and Statistical Psychology, 21,* 85-96.

Kaiser, H. F. (1958). The varimax criterion of analytic rotation in factor analysis. *Psychometrika, 2,* 187-200.

Kaiser, H. F., & Caffrey, J. (1965). Alpha factor analysis. *Psychometrika, 30, 1-14.*

Kaplan & Saccuzo (1982) *Psychological testing - Principles , applications and issues.* California: Cole Publishing Company.

Kausar, S., & Sheikh, H. (1986). Effect of Immediate vs. delayed knowledge of results on subsequent performance in verbal learning. *Pakistan Journal of Psychology, 17 (1-2)*, 21-32.

Keating, D. P. (1980). *Toward a multivariate developmental theory of intelligence. In D. Kuhn (ed.). Intellectual developmentbeyond chidhood.* San Francisco: Jossy.

Kelley, T. L. (1922). *The interpretation of educational measurements.* Yonkers-on-Hudson,

Kelley, T. L. (1928). *Crossroads in the mind of man: A study of differentiable mental abilities.* Stanford, California: Stanford University Press.

Khalid, H(1986). *Differences between primary students of government and private schools in cognitive development.* Unpublished M.

Khalique, N(1983). Effect of language similarity between the experimenter and the subject on the speed of verbal learning. *Pakistan Journal of Psychology, 14 (1-2)*, 17-22.

Khalique, N. (1981). Effect of physical presence of others on the speed of verbal learning of introverts and extroverts. *Pakistan Journal of Psychology, 12 (1-2),* 33-38.

Khalique, N. (1982). Audience effect and coaction effect on the speed of vocabulary writing. *Pakistan Journal of Psychology, 13 (3-4)*, 3-9.

Khan, S. (1987). Relationship between ideational fluency and intelligence among academically gifted and average students. *Pakistan Journal of Psychology, 18 (1-4)*, 17-28.

Kline, P. (1991). *Intelligence: The psychometric view.* New York: Routledge.

Krug, R.E. (1972) *Ability testing in developing countries, a hand book of principles and techniques.* New York: Praeger Publishers.

Kuder, G.F., & Richardson, M. W. (1937).The theory of estimation of test reliability. *Psychmetrica,2,*151-160.

Kuhn, D., Pennington, N., & Leadbeater, B. (1983).*Adult thinking in developmental perspective*. In P.B Baltes & O.G. Brim, Jr, (Eds). *Life development and behavior (Vol-5)*. New York: Academic Press.

Labouvie-Vief, G. (1980). Beyond formal operations: uses and limits of pure logic in life-span development. *Human Development, 23*, 141-161.

Labouvie-Vief, G. (1982).Dynamic development and mature autonomy: A theoretical prologue. *Human Development, 25*,161-191.

Lavin, D.E. (1965). *The prediction of academic performance.* New York: Willey.

Lenning, O.T. (1975). *Predictive validity of the ACT tests at selective colleges* (Report No.69) Lowa City : Research and Development

Levy, P. (1973). *On the relation between test theory and psychology.* In P. Kline (Ed.), *New approaches in psychological measurement.* London & New York: Wiley.

Linden, K. W., & Linden, J. D. (1968). *Modern mental measurement: A historical perspective.* Boston: Houghton-Mifflin.

Loevinger, J. (1947). A systematic approach to the construction and evaluation of test of ability. *Psychological Monographs, 61(4).*

Loevinger, J., Gleser, G. C., & DuBois, P. H. (1953). Maximizing the discriminating power of a multiple-score test. *Psychometrika, 18,* 309-317.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores.* Reading, Mass Addison-Wesley.

Lorge, I. (1945). Schooling makes a difference. *Teachers College Record, 46,* 483-492.

Lunzer, E. A. (1965). Problems of formal reasoning in test situation. *Monographs of the society for Research in Child Development. 30, (2),* 19-46.

Lynn, R. (1987). The intelligence of the mongoloids: A psychometric, evolutionary and neurological theory. *Personality and individual differences, 8,* 813-844.

Lynn, R., Pagliari, C., & Chann, J. (1988). Intelligence in Hong Kong measured for Spearman's g and the neurospatial and verbal primaries. *Intelligence, 12,* 423-433.

Mahmood, Z. (1991). Intelligence, IQ and the thrid world. *Pakistan Journal and Psychological Research,* Islamabad: National Institute of Psychology.

McClelland, D. C. (1993). Intelligence is not the best predictor of the job performance. *Current directions in Psychological Sciences, 2,* 5-6.

McClure, W. E. (1976). The status of psychological testing in large city public school systems. *Journal of Applied Psychology.* Chicago: Rand Mc Nally.

McDonald, R. P. (1962). A general approach to nonlinear factor analysis. *Psychometrika, 27*, 397-415.

Miles, T. K. (1957). Contributions to intelligence testing and the theory of intelligence. 1. On defining intelligence. *British Journal of Educational Psychology, 27,* 153-165.

Monroe, W. S. (1945). Educational measurement in 1920 and 1945. *Journal of Educational Research, 38,* 334-340.

Murphy, K.R., & Davideshafer, C.O. (1988). *Psycholigical testing - Principles and applications.* London :Prentice - Hall , International Inc.

Murphy, K.R., & Davideshafer, C.O. (1998). *Psycholigical testing - Principles and applications. (4th ed).* London :Prentice - Hall , International Inc.

Naheed, G (1993). *Development of verbal test of intelligence for Pakistani urban primary school children.* Unpublished M Phil Thesis. Islamabad: National Institute of Psychology.

Najam, N., Andrabi, N.,& Ghaznavi,M. (1990). Relationship between intelligence, order of birth and some personality variables. *Journal of Behavioural Sciences, 1 (2)*, 21-32.

Newmark, C. S. (1985). *Major Psychological Assessment Instruments.* London: Allyn and Bacon, Inc.

Novick, M. R. (1966). The axioms and principal results of classical test theory. *Journal of Mathematical Psychology, 3*,1-18.

Nunnally, J. C. (1970). *Introduction to Psychological Measurement*. New York: Mcgraw-hill book company.

Nunnaly, J. C. (1982). *Psychometric theory. (2nd ed)* NewYork: McGraw Hill.

Otis, A. S. (1918). An absolute point scale for the group measure of intelligence. *Journal of Educational Psychology, 9*, 238-261.

Owens, W. A. (1966). Age and mental abilities: A second adult follow-up. *Journal of Educational Psychology, 57*, 311-325.

Pearson, K. (1892). *The grammar of science.* London: Walter Scott.

Pearson, K. (1894). Contributions to the mathematical theory of evolution. I: On the dissection of asymmetrical frequency curves. *Philosophical Transactions (Ser. A), 185*, 71-110.

Pellegrino, J.W., & Glaser, R (1980) *Components of inductive reasoning.* In R.E. Snow, PA Federico, & W.E Montague (Eds). *Aptitude learning and instruction, cognitive process analyses of aptitude (Vol-I)* Hillsdale, New Jersy: Erbaum.

Pervez, M. (1982). *Cognitive developmental stages of primary school children of Pakistan: Phase I.* Islamabad: National Institute of Psychology, 47p.

Peterson, J. (1925). *Early conceptions and test of intelligence.* Yonkers-on-Hudson: World Book.

Piaget J. (1972) *The psychology of intelligence*. Totowa New Jersy: Littlefield, Adams.

Pintner, R. (1921). *Intelligence testing: Methods and results. (2nd ed.)*.NewYork: Henery Holt.

Rao, M, (1986). Validation of IEA Mathematics Achievement Test in Pakistan: A pilot study. *Pakistan Journal of Psychological Research, 1(3-4),* 51-55.

Rapaport, D., Gill, M. M., & Schafer, R. (1968). *Diagnostic psychological testing(Rev. ed.)(R. R. Holt.Ed.)*.New York: International universities press.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests.* Copehagen: Nielson & Lydiche.

Ree, M. J., Earles, J. A., & Teachout, M. S. (1994). Predicting training success : Not much more than *g*. *Journal of Applied Psychology, 79,*518-524.

Ree, M. J., & Carretta, T. R. (1994).The correlations of general cognitive ability and psychomoto traching tests. *International Journal of selection and assessment, 2,*209-216.

Ree, M. J., & Earles, J. A. (1991a). Predicting training success : Not much more than *g*. *Personnel Psychology, 44,* 321-332.

Ree, M. J., & Earles, J. A. (1991b). The stability of *g* : a method of estimation. *Intelligence, 15,*271-278.

Ree, M. J., & Earles, J. A. (1992). Intelligence is the best prediction of job performance. *Current directions in Psychological Sciences, 1,*86-89.

Reitman, W. (1965) *Cognition and thought*. New York: Wiley.

Riaz, M. N. (1979). A study of intelligence-creativity distinction and their relationship with academic achievement. *Pakistan Psychological Studies, 3,* 38-70.

Riaz, M. N. (1985). Creative potential and talented performance. New Directions in Pakistani Psychology: *Proceedings of the 5th Conference of Pakistan Psychological Associations.* Z. A. Ansari, P. N. Tariq, , Zaman, , R. M. Chowdhri, and S. Watanabe. (Editors). Islamabad: PPA, 39-43.

Riaz, M. N. (1989). Creativity and psycholigical differentiation in high and low achieving science students. *Pakistan Journal of Psychological Research, 4 (3-4),* 81-92.

Richardson, S. S. (1956). Some evidence relating to the validity of selection for grammar schools. *British Journal of Educational Psychology, 126,* 15-24.

Rogosa, D.A. (1980). critique of cross-lagged correlation, *Psychological Bulletin, 88,* 245-258.

Rumelhart, D. E., & Norman, D. A. (1981).*Analogical processes in learning. In J. R. Anderson (Ed). Cognitive skills and their acquisition.* New Jersy: Erlbaum.

Schmid, J., & Leiman. J. M. (1957). The development of hierarchical factor solutions. *Psychometricka, 22,* 53-61.

Sechrest, L., Fay, T. L., & Zaidi, S. M. H. (1972). Problems of translation in cross-cultural research. *Journal of Cross-cultural Psychology,3,* 41-56.

Shahid, S. (1989). *Development of a test for assessing acquisition of science concepts in secondary school students*. Unpublished M. Phil. Dissertation. Islamabad: National Institute of Psychology ,121 + 131p.

Sharp, S. E. (1899). Individual psychology: A study in psychological method. *American Journal of Psychology, 10,* 329-391.

Sheikh, M. H., & Raza, A. (1991).Translation and adaptation of Children's Personality Questionnaire (CPQ) of the IPAT series into Urdu and development of local norms for Pakistani population. *Pakistan Journal of Psychology, 22*, 25-37.

Skinner, B. F. (1953). *The science of human behavior*. New York: Macmillan.

Skyrms, B. (1975) *Choice and chance (2nd ed)*. Encino: CA Dickenson.

Snow, R.E. (1980). Aptitude and achievement. *New directions for testing and measurement, 5,* 39-59 (b).

Snow, R.E. (1981). *Toward a theory of aptitude for learning: I Fluid and crystallized abilities and their correlates.* In M. Friedman, J.P, das & N, O' Connor (Eds), Intelligence and learning New York: Plenum Press.

Solso, R. L. (1988). *Cognitive Psychology.2nd. edition*. University of Nevada-ramo. London: Allyn and Bacon, Inc.

Spearman, C. (1923). *The nature of "intelligence" and the principles of cognition.* London: Macmillan.

Spearman, C. (1927). *The abilities of man*. London: Macmillan.

Spearman, C., & Jones, L. W. (1950). *Human ability: A continuation of "The abilities of man."* London: Macmillan.

Stanley, J. C. (1970). Reliability. In R.L.Thorndike (1971). (Ed.), *Educational measurement (2nd ed.).* Washington, D. C.: American Council on Education.

Stanley, J. C., & Wang, M. D. (1970). Weighing test items and test-item options: An overview of the analytical and empirical literature. *Educational and Psychological Measurement, 30,* 21-35.

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

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

Sternberg, R. J. (Ed.) (1985c). *Human abilities: An information-processing approach.* London: W. H. Freeman.

Sternberg, R. J. (1986). Inside intelligence. *American Scientist, 74*, 137-143.

Sternberg, R. J., & Powell, J. S. (1963). Comprehending verbal comprehension. *American Psychologist, 38*, 878-893.

Sternberg, R. J. (1984). *Handbook of Intelligence.* New York:Cambridge University Press.

Sternberg, R,J., & Gardner, M.K. (1983). Unties in inductive reasoning. *Journal of experimental Psychology: General, 112*,80-116.

Sternberg, R,J., & Rifkin, B. (1979). The development of analogical reasoning processes. *Journal of Experimental Child Psychology, 27*,195-232.

Sternberg, R,J., Nigro,G (1980). Developmental patterns in the solution of verbal analogies. *Child Development, 51*,27-38.

Sternberg, R. J. (1977). *Intelligence, information processing and analogical reasoning: the componential analysis of human abilities.* Hillsdale, New Jersy: Erlbaum.

Sternberg, R. J. (1995).For whome the bell curve tallies: A review of the Bell Curve. *Psychological Science, 6,* 257-266.

Sternberg, R. J. (1997). Intelligence and life long learning. *American Psychologist,52,*1029-1139.

Sternberg, R. J., & Wagner, R. K. (1993).The *g-ocentric* view of intelligence and performance is wrong. *Current directions in Psychological Sciences, 2,*1-5.

Sternberg, R.J (1979) The nature of mental abilities *American Psychologist, 34*_214-230.

Sternberg, R.J (1984). Toward a trarchic theory of human intelligence. *Behavioral and Brain Sciences, 7,* 269-315.

Sternberg, R.J (Ed). (1982). *Handbook of human intelligence.* Cambridge: Cambridge University Press.

Sullivan, E. T., Clark, W. W., & Tegs, E. W. California(1939). *Test of Mental maturity.* Los Angeles: Californaia Test Bureau, 1936-1939.

Syed, N (1993). *Development of nonverbal test of intelligence for Pakistani urban primary school children.* Unpublished M Phil Thesis. Islamabad: National Institute of Psychology.

Terman, L. M. (1916). *The measurement of intelligence.* Boston: Houghton-Mifflin.

Thorndike, E. L. (1921). Intelligence and its measurement: A symposium. *Journal of Educational Psychology, 12*, 123-147, 195-216, 271-275.

Thorndike, R. L.& Hagen, E. (1959). *10,000 careers.* New York: Wiley.

Thorndike, E. L. (1904). *Introduction to the theory of mental and social measurements.* New York: Columbia University. Teachers College.

Thorndike, L. L. (1921). *The nature of intelligence.* NewYork: Harcourt.

Thorndike, R. L.& Hagen, E. P. (1977). *Measurement and evolution in psychology and education. (4th ed).* NewYork: Harcourt.

Thorndike, R.M., Cunningham, G.K., Thorndike, R.. L., & Hagen, E.P(1991). *Measurement and Evaluation in Psychology and Education, Fifth Edition.* New York : Macmillon Publishing Company.

Thurstone, L. L. (1925). A method of scaling psychological and educational test. *Journal of Educational Psychology, 16,* 433-451. (a)

Thurstone, L. L. (1926). Current issues in factor analysis. *Psychological Review, 33*, 268-278.

Thurstone, L.L., & Thurstone, T. G. (1962). *SRA Primary Mental Abilities.* Chicago: Science Research Associations, 1949-1965.

Tuddenham, R. D. (1962). *The nature and measurement of intelligence. In L. Postman (Ed.), Psychology in the making.* New York: Knopf.

Tyler, L. (1965). *The psychology of human differences,* New Jersy: Appleton-Century-Crofts.

Vernon, P. E. (1969). *Intelligence and cultural environment.* London: Methuen.

Vernon, P. E. (1979). *Intelligence, heredity and environment.* New York: W. H. Freeman.

Vernon, P. E. (1960). *Intelligence and attainment tests.* London: University of London Press.

Vernon, P. E. (1961). *The structure of human abilities (2nd ed.).* London: Methuen.

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

Warburton, F. W. (1951). The intelligence of the Gurkha recruit. *British Journal of Psychology, 42,* 123-133.

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

Wechsler,D. (1939). *The measurement of adult intelligence.* Baltimore:Williums and Wilkins.

Wechsler, D. (1981). *Mannual for Wechsler Ault Intelligence Scale Revised.* NewYork: Harcourt.

Wherry, R. J., Gaylord, R. H. (1943). The concept of test and item reliability in relation to factor pattern. *Psychometrika, 8,* 247-264.

Wherry, R. J., & Winer, B. J. (1953). A method for factoring large numbers of items. *Psychometrika,* 18, 161-179.

Whipple. G. M. (1910). *Mannual of mental and physical tests.* Baltimore: Warwick & York.

Williams, M. (1960). The effect of past experience on mental performance in the elderly. *British Journal of Medical Psychology, 33*, 215-219

Wolfle, D. (1940). Factor analysis . *Psychometric Monographs, 3.*

Wolman, B. B. (1982) *Hand Book of Intelligence - theories,*

Wundt, W. (1862). *Beitrage zur Theorie der Sinneswahrnehmung.* Leipzig & Heidelberg: Witner.

Zaidi, S. M. H. (1975). *Frontiers of Psychological Research in Pakistan. Karachi:* Psychology Department, Karachi University.

Zoofashan, T. (1982). *Computer model for statistical analysis of progressive matrices.* Unpublished M.Sc. Dissertation, Islamabad: Department of Computer Science, Quaid-i-Azam University.

ANNEXURES

Annexure-A

Sampling Plan of the candidates of PMA Long Course for present study

Study	City	Sample
First Experimental Tryout	Rawalpindi	200
Second Experimental Tryout	Rawalpindi	200
Time limit of the test	Rawalpindi	60
Reliability	Rawalpindi and Lahore	535
Validity	Rawalpindi and Lahore	535
Norm	Faisalabad	100
	Lahore	100
	Gujranwala	100
	Sialkot	100
	Multan	100
	Rawalpindi	100
	Quetta	100
	Peshawar	100
	Karachi	100
	Hyderabad	100

Annexture-B

Keys of the sub-tests of first try out

Answers	Items
	Vocabulary Test
1	2, 5, 8, 11, 14, 15, 18, 21, 24, 27, 29, 34, 39, 42, 46, 51, 54, 59, 63, 64, 67, 68, 69, 70, 73, 76, 79, 84, 87, 92, 94.
2	1, 6, 16, 23, 33, 36, 38, 41, 44, 49, 53, 58, 66, 74, 81, 83, 88, 89.
3	3, 7, 10, 12, 20, 22, 25, 26, 30, 32, 37, 40. 43, 47, 52, 55, 56, 60, 62, 65, 71, 75, 78, 80, 85, 91, 93, 95.
4	4, 9, 13, 17, 19, 28, 31, 35, 45, 48, 50, 57, 61, 72, 77, 82, 86, 90
	Verbal Reasoning Test
1	2, 18, 20, 25, 30, 34.
2	1, 4, 6, 10, 11, 12, 14, 16, 19, 22, 24, 26, 27, 29, 31, 33, 35, 36, 39, 40.
3	5, 7, 8, 9, 13, 15, 17, 21, 23, 28, 32, 37, 38.
4	3.
	Numerical Reasoning Test
1	10, 11, 12, 19, 23, 31, 32, 38, 43, 45, 51, 52, 54, 56.
2	5, 7, 15, 16, 18, 20, 21, 22, 26, 28,33, 40, 41, 44, 47, 48, 49, 50.
3	1, 6, 8, 13, 17, 24, 25, 27, 34, 35, 36, 37, 39, 42, 46, 55.
4	2, 3, 4, 9, 14, 29, 30, 53.
	Information Test
1	21, 22, 27, 31.
2	1, 3, 6, 8, 10, 12, 15, 19, 26, 28, 29, 32, 33, 35, 37, 39.
3	4, 5, 7, 9, 11, 14, 16, 17, 23, 25, 30, 34, 36, 38.
4	2, 13, 18, 20, 24, 40.

Annexure-C

Three indices obtained through item analyses of first experimental try out

Items	Difficulty level	Discrimination level	Internal Consistency
			Vocabulary Test
1.	94.53	-.02	.003
2.	46.09	.30	.24
3.	82.03	.30	.37
4.	75.78	.23	.18
5.	59.38	.28	.20
6.	89.84	.14	.23
7.	42.97	.33	.28
8.	10.94	.13	.17
9.	32.03	.17	.20
10.	76.56	.28	.27
11.	83.59	.33	.37
12.	85.16	.11	.12
13.	89.84	.17	.20
14.	61.72	.52	.39
15.	81.25	.28	.33
16.	32.03	-.02	-.01
17.	64.84	.33	.28
18.	49.22	.36	.26
19.	22.66	.11	.08
20.	30.47	.14	.18
21.	21.88	.16	.17
22.	56.25	.45	.34
23.	85.16	.14	.21
24.	92.19	.09	.18
25.	84.38	.22	.24

Cont....

26.	60.16	.48	.45
27.	72.66	.52	.48
28.	68.75	.41	.38
29.	53.13	.69	.50
30.	89.06	.16	.22
31.	53.91	.55	.38
32.	68.75	.44	.40
33.	65.63	.53	.45
34.	82.81	.13	.21
35.	86.72	.20	.30
36.	3.13	.06	.15
37.	50.00	.47	.37
38.	13.20	-.05	-.01
39.	53.13	.22	.22
40.	75.78	.30	.34
41.	67.97	.42	.38
42.	90.63	.16	.34
43.	75.78	.45	.49
44.	39.84	.48	.40
45.	58.59	.30	.26
46.	46.09	.42	.30
47.	44.53	.64	.48
48.	57.81	.44	.39
49.	64.06	.44	.35
50.	16.41	.17	.21
51.	76.56	.47	.49
52.	15.63	.03	.03
53.	84.38	.31	.47
54.	47.66	.42	.35
55.	53.13	.44	.35

Cont...

56.	81.25	.25	.35
57.	32.81	.50	.41
58.	4.69	-.03	-.09
59.	16.41	.05	.07
60.	39.06	.53	.41
61.	28.91	.02	.01
62.	62.50	.56	.47
63.	34.38	.03	.07
64.	21.88	.13	.13
65.	64.84	.48	.42
66.	17.19	.16	.19
67.	41.41	.33	.30
68.	72.66	.23	.26
69.	84.38	.28	.39
70.	28.13	.31	.31
71.	68.75	.53	.48
72.	49.22	.67	.46
73.	47.66	.45	.36
74.	38.28	.39	.27
75.	81.25	.34	.45
76.	21.88	.34	.34
77.	58.59	.30	.25
78.	54.69	.72	.53
79.	32.03	.33	.27
80.	59.38	.66	.52
81.	50.78	.55	.43
82.	68.75	.44	.41
83.	35.16	.23	.24
84.	40.63	.50	.39
85.	46.88	.56	.43

Cont...

86.	50.00	.22	.17
87.	48.44	.44	.36
88.	62.50	.50	.42
89.	43.75	.56	.45
90.	18.75	.19	.21
91	77.34	.33	.36
92.	53.13	.59	.42
93.	35.16	.39	.29
94.	31.25	.22	.21
95.	22.66	.23	.22

Verbal Reasoning Test

1.	70.31	.41	.40
2.	85.94	.28	.70
3.	78.13	.41	.53
4.	77.34	.39	.62
5.	85.94	.28	.68
6.	82.03	.36	.65
7.	73.44	.34	.39
8.	85.94	.28	.65
9.	85.16	.27	.69
10	89.84	.20	.73
11.	80.47	.36	.57
12.	85.94	.25	.70
13.	81.25	.31	.48
14.	78.91	.39	.54
15.	23.44	.25	.17
16.	78.91	.42	.67
17.	85.94	.25	.66
18.	81.25	.38	.75

Cont...

19.	84.38	.31	.66
20.	75.00	.50	.60
21.	60.16	.52	.48
22.	77.34	.39	.66
23.	70.31	.38	.46
24.	78.13	.41	.66
25.	84.38	.31	.73
26.	25.78	.20	.18
27.	77.34	.39	.62
28.	82.81	.31	.74
29	56.25	.47	.33
30.	40.63	.47	.33
31.	78.91	.42	.70
32.	79.69	.41	.78
33.	71.88	.53	.68
34.	78.13	.44	.72
35.	69.13	.55	.51
36.	68.75	.53	.63
37.	80.47	.39	.82
38.	77.34	.42	.68
39.	66.41	.55	.57
40.	80.47	.39	.70

Numerical Reasoning Test

1.	82.03	.36	.57
2.	55.47	.27	.24
3.	82.81	.31	.46
4.	41.44	.38	.28
5.	80.47	.30	.49
6.	76.56	.34	.44

7.	63.28	.58	.53
8.	62.50	.53	.47
9.	77.34	.39	.57
10	56.25	.41	.36
11.	71.88	.44	.47
12.	36.72	.30	.26
13.	39.84	.33	.24
14.	48.44	.16	.20
15.	37.50	.25	.21
16.	76.56	.38	.47
17.	59.38	.53	.44
18.	62.50	.50	.43
19.	58.59	.52	.42
20.	53.91	.61	.45
21.	39.06	.28	.26
22.	72.66	.36	.37
23.	69.53	.48	.52
24.	75.00	.44	.50
25.	63.28	.48	.53
26.	57.81	.63	.46
27.	73.44	.41	.39
28.	49.22	.70	.51
29	71.88	.44	.55
30.	77.34	.42	.55
31.	75.78	.36	.47
32.	17.97	.20	.19
33.	76.56	.44	.60
34.	55.47	.61	.53
35.	57.07	.64	.46
36.	60.16	.48	.41

Cont...

37.	72.66	.52	.55
38.	71.88	.44	.51
39.	54.69	.69	.47
40.	64.84	.55	.50
41.	62.50	.66	.56
42.	62.50	.63	.59
43.	71.09	.52	.60
44.	69.53	.52	.56
45.	57.81	.59	.44
46.	28.13	.44	.35
47.	61.72	.61	.53
48	35.16	.39	.30
49.	68.75	.56	.60
50.	46.88	.63	.51
51.	38.28	.45	.37
52.	43.75	.47	.37
53.	66.41	.58	.51
54.	49.22	.70	.52
55.	50.00	.59	.46
56.	25.00	.34	.28

Information Test

1.	84.38	.28	.62
2.	67.19	.44	.45
3.	39.84	.33	.30
4.	42.97	.36	.29
5.	25.00	.13	.15
6.	57.03	.39	.42
7.	75.78	.33	.44
8.	57.81	.56	.46

Cont...

9.	46.88	.34	.26
10	46.09	.27	.28
11.	59.53	.55	.53
12.	67.97	.42	.43
13.	08.59	.08	.10
14.	64.84	.58	.56
15.	80.47	.39	.76
16.	50.00	.41	.36
17.	75.00	.44	.60
18.	55.47	.70	.54
19.	67.19	.59	.59
20.	66.41	.55	.58
21.	64.84	.52	.55
22.	65.63	.59	.63
23.	59.38	.66	.58
24.	60.94	.72	.67
25.	60.94	.63	.57
26.	63.28	.55	.61
27.	66.41	.58	.60
28.	76.56	.47	.74
29	37.50	.41	.32
30.	49.22	.48	.39
31.	50.78	.52	.44
32.	56.63	.53	.60
33.	49.22	.52	.46
34.	62.50	.66	.57
35.	46.09	.58	.49
36.	36.72	.61	.41
37.	60.16	.70	.56
38.	25.00	.25	.23

Cont...

39.	35.94	.44	.36
40.	50.78	.67	.50

Annexture-D

Keys of the sub-tests of second try out

Answers	Items	Tests
		Vocabulary Test
1	6,10, 11, 18, 22, 26, 30, 37, 38, 45, 46, 53, 54	
2	1, 5, 7, 12, 17, 19, 23, 27, 31, 36, 39, 44, 47, 52.	
3	2, 4, 8, 13, 16, 20, 24, 28, 32, 35, 40, 43, 48, 51.	
4	3, 9, 14, 15, 21, 25, 29, 33, 34, 41, 42, 49, 50.	
		Verbal Reasoning Test
1	2, 7, 8, 15, 16, 23, 30, 31.	
2	1 , 4, 6, 9, 17, 22, 24, 29, 32.	
3	5, 10, 13, 14, 18, 21, 25, 27, 28, 33.	
4	3, 11, 12, 19, 20, 26.	
		Numerical Reasoning Test
1	1, 7, 8, 11, 13, 21, 22, 27, 35, 41, 42.	
2	2, 4, 10, 12, 14, 18, 19, 20, 23, 26, 28, 33, 34, 36.	
3	3, 5, 9, 15, 17, 24, 25, 29, 30, 32, 37, 38, 40.	
4	6, 16, 31, 39.	
		Information Test
1	2, 3, 4, 29.	
2	1, 5, 7, 8, 9, 11, 14, 17, 18, 20, 25, 27, 28, 30, 31, 34, 36.	
3	6, 10, 12, 13, 16, 19, 21, 23, 24, 32, 35.	
4	15, 22, 26, 33.	

Annexure-E

Three indices obtained through item analyses of second experimental try out

Items	Difficulty level	Discrimination level	Internal Consistency
			Vocabulary Test
1	87.96	.17	.17
2	37.96	.31	.32
3	64.81	.56	.44
4	39.81	.69	.55
5	64.81	.44	.38
6.	96.30	.07	.19
7.	95.37	.09	.21
8.	21.30	.24	.32
9.	52.78	.72	.50
10.	89.81	.13	.16
11.	65.74	.43	.39
12.	66.67	.37	.28
13.	33.33	.48	.41
14.	55.56	.30	.24
15.	75.00	.39	.38
16.	66.67	.63	.47
17.	74.07	.41	.34
18.	75.93	.33	.33
19.	44.44	.59	.46
20.	82.41	.28	.42
21.	62.04	.39	.36
22.	55.56	.70	.54
23.	64.81	.63	.52
24.	33.33	.30	.24

Cont....

25.	14.81	.07	.08
26.	73.15	.46	.43
27.	80.56	.31	.34
28.	43.52	.76	.61
29.	75.00	.39	.36
30.	65.74	.46	.38
31.	62.96	.19	.20
32.	62.96	.56	.46
33.	52.78	.80	.58
34.	28.70	.28	.21
35.	59.26	.67	.51
36.	46.30	.33	.32
37.	63.89	.24	.23
38.	46.30	.41	.35
39.	65.74	.28	.25
40.	68.52	.33	.30
41.	59.26	.44	.37
42.	21.30	.20	.25
43.	57.41	.48	.39
44.	20.37	.19	.24
45.	54.63	.57	.45
46.	45.37	.28	.25
47.	29.63	.33	.34
48.	65.74	.24	.20
49.	31.48	.44	.44
50.	53.70	.63	.44
51.	32.41	.28	.26
52.	62.96	.44	.36
53.	31.48	.26	.25
54.	51.85	.30	.24

Cont....

			Verbal Reasoning Test
1.	64.81	.37	.36
2.	100.00	.00	00
3.	80.56	.35	.45
4.	95.37	.09	.44
5.	97.22	.06	.44
6.	94.44	.07	.29
7.	89.81	.20	.40
8.	89.81	.20	.50
9.	89.81	.20	.57
10.	84.26	.31	.44
11.	96.30	.07	.40
12.	75.00	.46	.40
13.	76.85	.43	.42
14.	22.22	.26	.20
15.	93.52	.09	.41
16.	93.52	.13	.47
17.	92.59	.15	.42
18.	93.52	.13	.52
19.	81.48	.37	.34
20.	63.89	.61	.51
21.	86.11	.24	.39
22.	60.19	.39	.33
23.	89.81	.20	.60
24.	91.67	.17	.70
25.	87.96	.20	.56
26.	87.04	.26	.59
27.	75.00	.31	.29
28.	92.59	.15	.63
29.	95.37	.09	.52

Cont....

30.	89.81	.20	.51
31	79.63	.33	.55
32.	86.11	.20	.29
33.	89.81	.20	.52
			Numerical Reasoning Test
1.	92.52	.09	.56
2.	88.89	.15	.44
3.	75.00	.31	.44
4.	64.81	.63	.55
5.	85.19	.19	.42
6.	86.11	.28	.49
7.	71.30	.35	.41
8.	47.22	.39	.29
9.	69.44	.39	.42
10.	65.74	.35	.33
11.	36.11	.35	.29
12.	51.85	.52	.40
13.	78.70	.28	.38
14.	35.70	.44	.37
15.	43.52	.46	.34
16.	76.85	.39	.48
17.	86.11	.24	.48
18.	61.11	.44	.41
19.	56.48	.46	.39
20.	79.63	.33	.44
21.	87.96	.20	.45
22.	85.19	.26	.53
23.	34.26	.24	.18
24.	73.15	.46	.52

Cont....

25.	64.81	.52	.41
26.	62.96	.48	.42
27.	67.59	.50	.50
28.	76.85	.43	.50
29.	77.78	.37	.61
30.	85.19	.30	.54
31.	54.63	.65	.52
32.	42.59	.67	.43
33.	75.00	.46	.49
34.	43.52	.54	.38
35.	80.56	.28	.53
36.	00	00	.46
37.	67.59	.35	.40
38.	47.22.	.57	.39
39.	81.48	.37	.57
40.	61.11	.44	.36
41.	8.33	.02	.04
42.	50.93	.54	.42
			Information Test
1.	90.74	.19	.50
2.	92.59	.15	.51
3.	26.04	.17	.21
4.	37.04	.11	.13.
5.	66.67	.37	.35
6.	70.37	.52	.54
7.	65.74	.39	.39
8.	49.07	.39	.30
9.	42.59	.19	.17
10.	74.07	.48	.56
11.	66.67	.56	.44

Cont....

12.	61.11	.41	.36
13.	78.70	.35	.54
14.	84.26	.31	.67
15.	48.15	.48	.42
16.	77.78	.44	.64
17.	66.67	.52	.52
18.	73.15	.31	.49
19.	54.63	.61	.51
20.	64.81	.44	.52
21.	71.30	.50	.58
22.	62.04	.61	.48
23.	75.93	.48	.66
24.	67.59	.46	.51
25.	73.15	.50	.63
26.	81.48	.37	.70
27.	40.74	.37	.30
28.	53.70	.56	.47
29.	58.33	.50	.40
30.	70.37	.41	.52
31.	72.22	.56	.70
32.	60.19	.57	.51
33.	63.89	.72	.60
34.	32.41	.39	.29
35.	62.96	.63	.53
36.	38.89	.22	.25

Annexture-F

Keys of the sub-tests of SVITU (Final Draft)

Answers	Items
	Vocabulary Test
1	7, 8, 12, 19, 23, 31, 32, 38, 41.
2	1, 6, 9, 13, 18, 22, 26, 30, 33, 37, 40, 42.
3	2, 5, 10, 14, 17, 21, 25, 29, 34, 36, 39.
4	3, 4, 11, 15, 16, 20, 24, 27, 28, 35.
	Verbal Reasoning Test
1	1, 5, 11, 17.
2	2, 6, 10, 12, 16, 19, 20.
3	3, 7, 9, 13, 15, 18.
4	4, 8, 14.
	Numerical Reasoning Test
1	4, 5, 9, 16, 17, 24, 25, 32, 33.
2	3, 6, 10, 15, 18, 23, 26, 31, 34.
3	2, 7, 11, 14, 19, 22, 27, 28, 30, 35.
4	1, 8, 12, 13, 20, 21, 29, 36.
	Information test
1	1,7,13, 19, 25.
2	2, 6, 8, 12, 14, 18, 24, 26, 30.
3	3, 5, 9, 11, 15, 17, 20, 21, 23, 27, 29.
4	4, 10, 16, 22, 28.

Annexture-F

Keys of the sub-tests of HT (Final Draft)

Answers	Items	Tests
		Vocabulary Test
1	7, 8, 12, 19, 23, 31, 32, 38, 41.	
2	1, 6, 9, 13, 18, 22, 26, 30, 33, 37, 40, 42.	
3	2, 5, 10, 14, 17, 21, 25, 29, 34, 36, 39.	
4	3, 4, 11, 15, 16, 20, 24, 27, 28, 35.	
		Verbal Reasoning Test
1	1, 5, 11, 17.	
2	2, 6, 10, 12, 16, 19, 20.	
3	3, 7, 9, 13, 15, 18.	
4	4, 8, 14.	
		Numerical Reasoning Test
1	4, 5, 9, 16, 17, 24, 25, 32, 33.	
2	3, 6, 10, 15, 18, 23, 26, 31, 34.	
3	2, 7, 11, 14, 19, 22, 27, 28, 30, 35.	
4	1, 8, 12, 13, 20, 21, 29, 36.	
		Information test
1	1,7,13, 19, 25.	
2	2, 6, 8, 12, 14, 18, 24, 26, 30.	
3	3, 5, 9, 11, 15, 17, 20, 21, 23, 27, 29.	
4	4, 10, 16, 22, 28.	

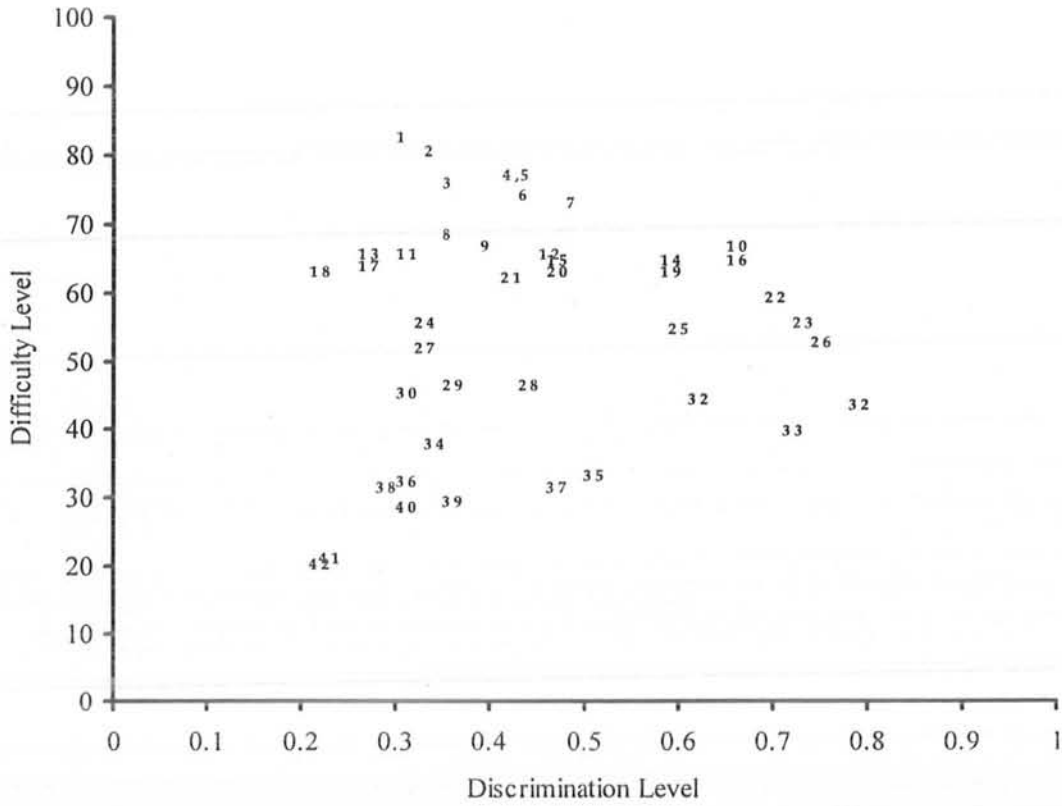

Figure 1: Items Dispersion of Vocabulary Test (Final Draft)

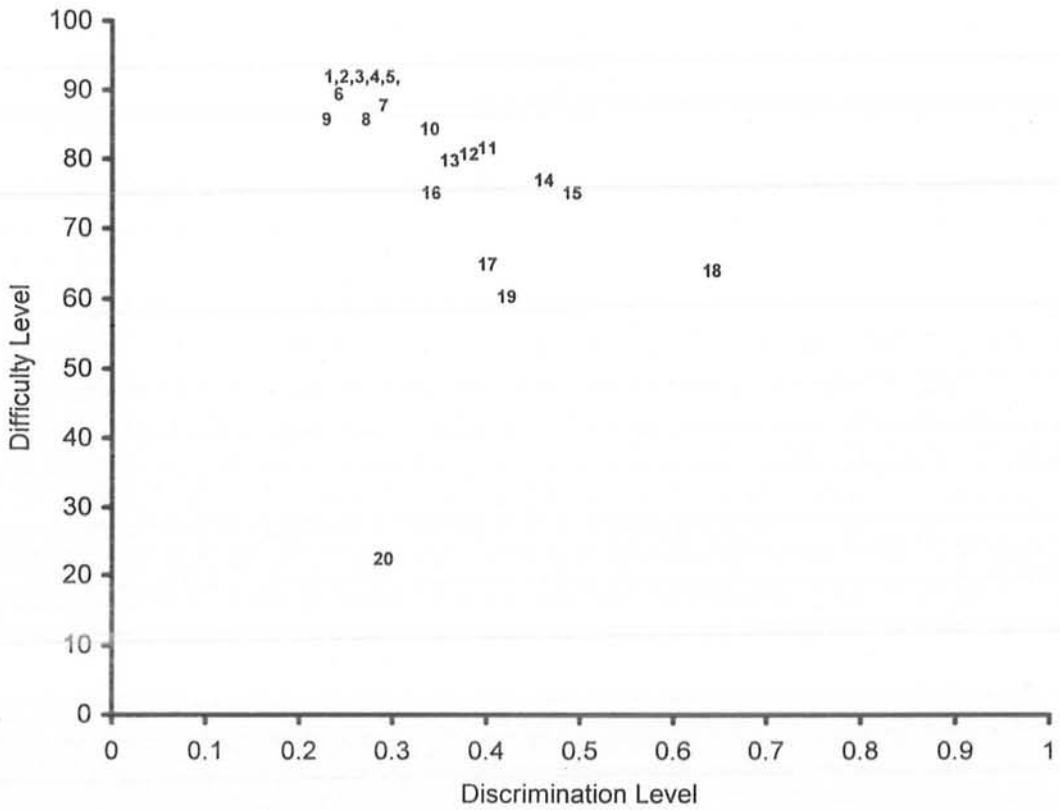

Figure 2: Items Dispersion of Verbal Reasoning Test (Final Draft)

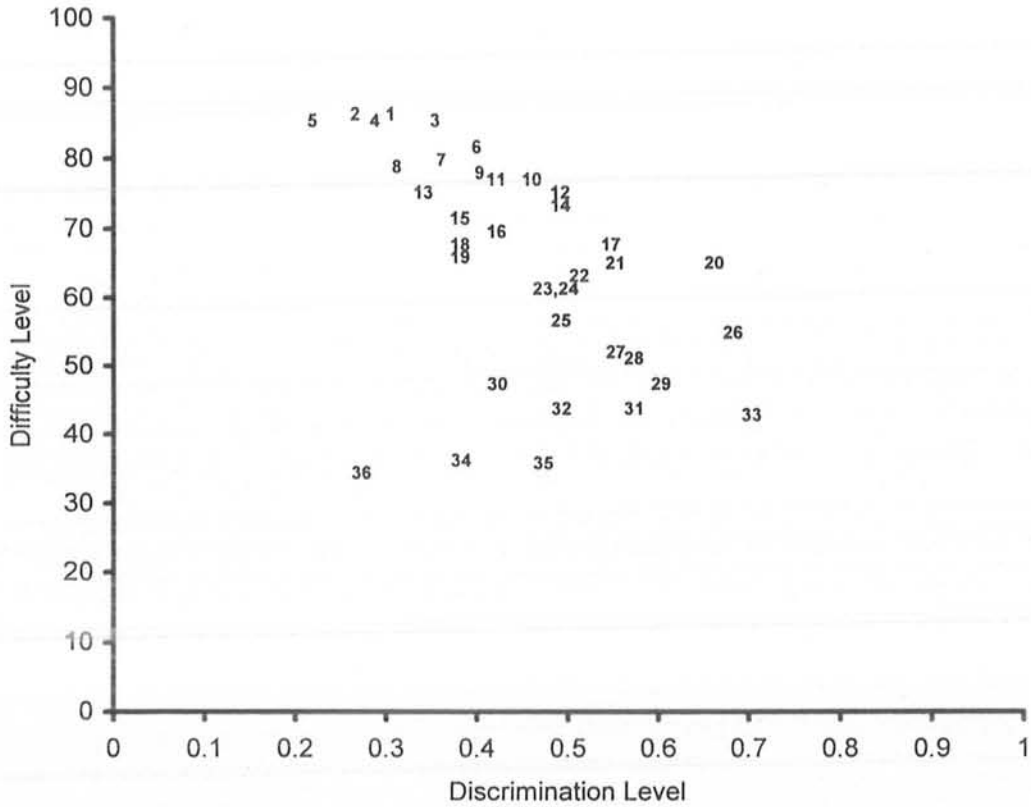

Figure 3: Items Dispersion of Numerical Reasoning Test (Final Draft)

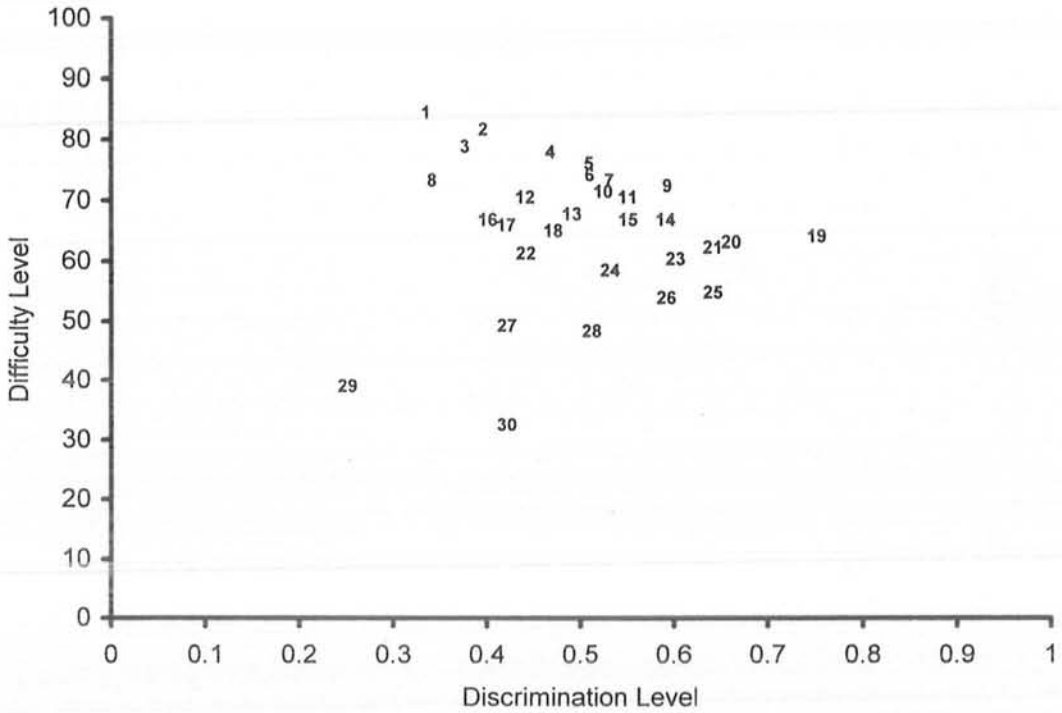

Figure 4: Items Dispersion of Information Test (Final Draft)

ذہانت کی آزمائش

ہدایات

عزیز بچو! آپ ذہانت کی آزمائش میں شریک ہوں گے۔اس آزمائش کے چار حصے ہیں۔حصہ اوّل الفاظ کے معانی پر مشتمل ہے،حصہ دوئم زبانی استدلال پر مشتمل ہے،حصہ سوئم عددی استدلال پر مشتمل ہے اور حصہ چہارم معلومات پر مشتمل ہے۔

آپ ذہانت کی آزمائش کے ہر حصے کو حل کرنے سے قبل اس حصے میں دی گئی متعلقہ ہدایات غور سے پڑھیں اور ان پر سختی سے عمل کریں۔

ذہانت کی آزمائش کے جوابات دینے کے لئے آپ کو ایک عدد جوابی فارم دیا جا رہا ہے۔ آپ ذہانت کی آزمائش کے ہر حصے سے سوالات پڑھیں اور جوابی فارم پر ہر سوال کے متعلقہ جوابی خانہ (ج) میں صرف ایک علامت (ا ، ب ، ج یا د) درج کریں۔

نمونہ برائے سوالات

نمبر شمار	سوالات	ا	ب	ج	د
.1	عافیت	آخرت	خیریت	واقفیت	مہربانی
.2	چڑیا کا گھونسلے سے وہی تعلق ہے جو چوہے کا تعلق ---- سے ہے۔	بل	چوہیا	بل	گھر
.3	اگلا عدد بتائیں 8, 6, 6, 4, 4, 2, 2.....	8	7	6	5
.4	پاکستان کے قومی پرچم میں کتنے رنگ ہیں؟	4	1	3	2

نمونہ برائے جوابی فارم

س	ج	س	ج	س	ج	س	ج
1	ب	2	ج	3	ا	4	د

حصّہ اوّل

الفاظ کے معانی

☆ **ہدایات** : عزیز بچو! اس ٹیسٹ میں آپ پہلے ہر لفظ غور سے پڑھیں۔ ہر لفظ کے سامنے جوابات میں چار معنی دیئے گئے ہیں۔ آپ صرف درست معنی کو چنیں اور جوابی فارم پر ہر لفظ کے متعلقہ نمبر کے سامنے اس کے جوابی خانہ میں صرف ایک علامت (ا ، ب ، ج یا د) درج کریں۔ یاد رکھیں ہر لفظ کا درست معنی صرف ایک ہی ہے۔

اس حصّے کے لئے مقررہ وقت 12 منٹ ہے۔

نمبر شمار	الفاظ	ا	ب	ج	د
1	مَتروک	مشترک	ترک کیا ہوا	پاک	مبارک
2	تلافی	معافی	تلف کرنا	نقصان کا عوض	فدیہ
3	وَصف	طریقہ	درمیان	وقف	خوبی
4	محقق	حقدار	پرہیزگار	لکھنے والا	تحقیق کرنے والا
5	تغافُل	لاپتہ	فکر	بے پرواہی	اعتبار
6	طِفل	چاند	بچہ	تالا	چھوٹا
7	مصوّر	تصور کرنے والا	مطلوب	تصویر بنانے والا	تخلیق کار
8	مُتغیّر	تبدیل شدہ	غیر	متبادل	بے ترتیب
9	حُرمَت	حرکت	عزت	تکلیف	بے عزتی
10	دائمی	پرانی	عارضی	ابدی	دوسری
11	تحقیر	جستجو	مذمت	کم ظرف	حقارت
12	مُحرّک	تحریک کرنے والا	وجہ	مہلک	معلق
13	رقابت	ملاپ	محبت میں شرکت	رفاقت	بندگی
14	داروغہ	جلاد	دعاگو	محافظ	قصہ گو
15	وُثوق	ذریعہ	جلدی	ورثہ	اعتماد
16	نامُوس	ختم	ناخوش	مایوس	عزت

نمبر شمار	الفاظ	ا	ب	ج	د
17	مانع	مضر	مطلب	منع کرنے والا	مانند
18	سہڈول	متواتر	خوبصورت	نڈھال	زخمی
19	دقیق	مشکل	دوست	حقارت	کم
20	منع	تنبیہہ	طبیعت	نصیحت	نکلنے کی جگہ
21	پھبتی	تحریک	پاگل	طنز	الزام
22	بےاعتنائی	بے گناہی	بے پرواہی	اجنبی	بے صبری
23	مجاور	خادم	جامع	مجازی	جادو
24	نمود	پرورش	صبح	تباہ	نمائش
25	گزند	راہ گیر	گزارنا	نقصان	غلطی
26	سموم	معصوم	گرم ہوا	موسم	سونا
27	خرد	درخت	خوراک	دھوکا	عقل
28	فریفتہ	زمہ داری	دھوکے باز	فریضہ	عاشق
29	معدوم	معذور	اداس	ناپید	شرمندہ
30	نشاط	نشست	خوشی	حیرت	نشہ
31	صریح	صاف	لمبا قد	معمولی	مشکل
32	بلوہ	ہنگامہ	بلانا	نظارہ	فتح
33	نخل	ویرانہ	درخت	نخرہ	ریگستان
34	قباحت	نفرت	اکتانا	برائی	بےوقوفی
35	مدفن	دفن شدہ	منصب	تعارف	بدبودار
36	خفیف	چھپا ہوا	ڈر	تھوڑا	کنجوس
37	حقیر	کم ظرف	چھوٹا	کمزور	عاجز
38	ڈھنڈورا	تشہیر	تلاش	شور	قاصد
39	استقلال	صبر	بہادری	مضبوطی	میانہ روی
40	ادراک	موقع	فہم	خاتمہ	بچنا
41	فروتنی	عاجزی	فراوانی	ذیادتی	ضروری
42	ہویدا	راز	واضح	ہوادار	حیران

زبانی استدلال

ہدایات : عزیز بچو! اس ٹیسٹ میں شامل ہر جملے کو غور سے پڑھیں۔ ہر جملے میں موجود دو جوڑوں کو اسطرح نامکمل رکھا گیا ہے تاکہ آپ کے دیے گئے اب کا پہلا حصہ پہلے جوڑے کے پہلے حصے کو مکمل کرے اور دوسرا حصہ دوسرے جوڑے کے آخری حصے کو مکمل کرے۔ آپ کا جواب ا،ب،ج،یا د ے صرف ایک علامت ہونا چاہیے۔ جو آپ جوابی فارم پر جوابی خانے میں درج کریں گے۔

اس حصے کے لئے مقررہ وقت 6 منٹ ہے۔

نمبر شمار	جملے	ا	ب	ج	د
1۔	---- کا پیچھے سے وہی تعلق ہے جو اوپر کا تعلق ---- سے ہے۔	آگے ۔ نیچے	ہار ۔ جیت	آہستہ ۔ تیز	دائیں ۔ بائیں
2۔	---- کا پھیپھڑے سے وہی تعلق ہے جو دل کا تعلق ---- سے ہے۔	ہوا ۔ پانی	ہوا ۔ خون	جسم ۔ جان	ہوا ۔ دماغ
3۔	---- کا بیڈ سے وہی تعلق ہے جو کرسی کا تعلق ---- سے ہے۔	ہسپتال۔ سکول	ہسپتال ۔ سٹیشن	ہسپتال ۔ ڈاکٹر	ڈاکٹر ۔ استاد
4۔	---- کا انسان سے وہی تعلق ہے جو گھر کا تعلق ---- سے ہے۔	پاؤں ۔ جوتا	پاؤں ۔ گھوڑا	کام ۔ لوہار	لوہا ۔ گھوڑا
5۔	---- کا دریا سے وہی تعلق ہے جو ساحل کا تعلق ---- سے ہے۔	سیلاب ۔ ریت	کنارہ ۔ سمندر	موج ۔ سمندر	موج ۔ کنارہ
6۔	---- کا جیل سے وہی تعلق ہے جو ہسپتال کا تعلق ---- سے ہے۔	قید ۔ مریض	قیدی ۔ مریض	مجرم ۔ ڈاکٹر	قید ۔ وارڈ
7۔	---- کا مستقبل سے وہی تعلق ہے جو پچھتاوے کا تعلق ---- سے ہے۔	آگے ۔ پیچھے	امید ۔ اندازہ	امید ۔ ماضی	آئندہ ۔ گزشتہ
8۔	---- کا ڈاکٹر سے وہی تعلق ہے جو ملزم کا تعلق ---- سے ہے۔	نرس ۔ وکیل	مریض ۔ عدالت	نرس ۔ عدالت	مریض ۔ وکیل
9۔	---- کا سائنس سے وہی تعلق ہے جو صادقین کا تعلق ---- سے ہے۔	عبدالسلام ۔ شاعری	اقبال ۔ مصوری	عبدالسلام ۔ مصوری	اقبال ۔ شاعری
10۔	---- کا ریل گاڑی سے وہی تعلق ہے جو بندرگاہ کا تعلق ---- سے ہے۔	انجن ۔ بحری جہاز	سٹیشن ۔ ساحل	انجن ۔ ساحل	سٹیشن ۔ بحری جہاز
11۔	---- کا فاصلے سے وہی تعلق ہے جو گرام کا تعلق ---- سے ہے۔	میٹر ۔ بجلی	منزل ۔ وزن	میٹر ۔ وزن	منزل ۔ راستہ
12۔	---- کا جملے سے وہی تعلق ہے جو جملے کا تعلق ---- سے ہے۔	لفظ ۔ پیرا	لفظ ۔ جملہ	پیرا ۔ لفظ	لفظ ۔ حرف
13۔	---- کا گائے سے وہی تعلق ہے جو مرغی کا تعلق ---- سے ہے۔	گوشت ۔ انڈہ	دودھ ۔ مرغا	دودھ ۔ انڈہ	بیل ۔ انڈہ
14۔	---- کا جھیل سے وہی تعلق ہے جو جنگل کا تعلق ---- سے ہے۔	جانور ۔ گھاس	پانی ۔ جانور	بیلچ ۔ درخت	پانی ۔ درخت
15۔	---- کا زنجیر سے وہی تعلق ہے جو موتی کا تعلق ---- سے ہے۔	کڑی ۔ ہار	پہاڑ ۔ ہار	لوہا ۔ کڑی	گھڑی ۔ ہار
16۔	---- کا آواز سے وہی تعلق ہے جو اندھے کا تعلق ---- سے ہے۔	کان ۔ نظر	بہرہ ۔ گونگا	کان ۔ اندھیرہ	بہرہ ۔ نظر
17۔	---- کا بارہ سے وہی تعلق ہے جو دو کا تعلق ---- سے ہے۔	چودہ ۔ ایک	نو ۔ تین	نو ۔ چھ	چار ۔ چھ
18۔	---- کا حال سے وہی تعلق ہے جو ادھار کا تعلق ---- سے ہے۔	مستقبل۔ ماضی	مستقبل۔ حال	ماضی۔ نقد	مستقبل۔ نقد
19۔	---- کا دن سے وہی تعلق ہے جو کیلینڈر کا تعلق ---- سے ہے۔	گھڑی ۔ سال	رات ۔ سال	چاند ۔ سال	سورج ۔ سال
20۔	---- کا چوٹی سے وہی تعلق ہے جو لہر کا تعلق ---- سے ہے۔	پہاڑ ۔ گہرائی	اوپر ۔ بل	اونچائی ۔ گہرائی	اوپر ۔ پانی

حصّہ سوئم

عددی اِستدلال

☆ **ہدایات** : عزیز بچو! آپ پہلے ہر سوال غور سے پڑھیں۔ ہر سوال کے چار جوابات دیئے گئے ہیں۔ آپ جوابات میں سے صرف درست جواب کو چنیں اور جوابی فارم پر ہر سوال کے متعلقہ نمبر کے سامنے اس کے جوابی خانہ میں صرف ایک علامت (ا ، ب ، ج یا د) درج کریں۔ یاد رکھیں ہر سوال کا درست جواب صرف ایک ہی ہے۔

اس حصّے کے لئے مقررہ وقت 16 منٹ ہے۔

نمبر شمار	سوالات	ا	ب	ج	د
.1	3 سال پہلے اکبر 5 سال کا تھا اور اس کی بہن 8 سال کی تھی۔ اکبر کی موجودہ عمر کتنے سال ہے؟	5	6	7	8
.2	اگلا عدد بتائیں 2, 8, 4, 8, 6, 8, ...	6	7	8	9
.3	اگر ایک انڈہ 5 منٹ میں ابلے تو اکٹھے 5 انڈے کتنے منٹ میں ابلیں گے؟	10	5	20	25
.4	اگلا عدد بتائیں 66, 70, 74, 78, ...	82	84	80	86
.5	اگلا عدد بتائیں 4, 8, 16, 32, ...	64	63	62	60
.6	اگلا عدد بتائیں 1, 2, 2, 3, 3, 4, ...	3	4	5	6
.7	اگر احمد، علی سے آہستہ چلے، طاہر، علی کے برابر چلے تو کون سب سے آہستہ چلے گا؟	طاہر	علی	احمد	کوئی نہیں
.8	کس عدد کو 12 میں جمع کیا جائے کہ 9 کا دگنا ہو جائے؟	2	4	8	6
.9	اگلا جوڑا مکمل کریں (20,10) (25,10) (30,10,....)	30	35	40	45
.10	اگلا جوڑا مکمل کریں (3,3) (9,3) (3 , ...)	21	27	25	23
.11	حل بتائیں 2 سال 8 ماہ 30 دن + 3 سال 3 ماہ 10 دن = ...	4سال	5سال	6سال	7سال
.12	اگلا جوڑا مکمل کریں (21,3) (42,6) (9 ,....)	21	72	42	63
.13	احمد 44 سال کی عمر میں موجودہ عمر سے 4 گنا زیادہ ہوگا تو احمد کی موجودہ عمر کتنے سال ہوگی؟	40	33	22	11
.14	حل بتائیں 2 -- 2 X 2 + 2 ÷ 2 = ...	0	1	2	4

نمبر شمار	سوالات	ا	ب	ج	د
.15	اگر بابر، اسلم سے چھوٹا ہو، احمد، بابر سے بڑا ہو اور اسلم، احمد سے چھوٹا ہے۔ تو سب سے بڑا کون ہوگا؟	بابر	احمد	اسلم	کوئی نہیں
.16	24 کا 25% کیا ہوگا؟	6	12	24	25
.17	اگلا عدد بتائیں 6 7 8 , 6 8 7 , 8 6	7	5	9	3
.18	اگلا عدد بتائیں 8, 6, 7, 5, 4	6	9	8	7
.19	اگر ایک ڈرائیور 120 کلو میٹر فی گھنٹہ کی رفتار سے گاڑی چلائے تو 30 کلو میٹر کا فاصلہ کتنے منٹ میں طے کرے گا؟	10	30	15	35
.20	اگر گزرے ہوئے پرسوں 9 تاریخ تھی تو آنے والے پرسوں کونسی تاریخ ہوگی؟	9	11	15	13
.21	اگلا عدد بتائیں ....., 3, 4, 2, 5, 1, 6	4	1	2	3
.22	حل بتائیں ..... = 3/4 x 4/3	0	3	1	4
.23	حل بتائیں ..... = .05 x .5	.05	.025	.25	.0025
.24	مساوات مکمل کریں ..../77 = 11/5	35	30	7	5
.25	حل بتائیں ..... = 2/4 + 3/6	1	2	3	4
.26	اگر ا = 2 ، د = 3 ، ر = 4 اور ک =5 تو لفظ ا د ر اک میں شامل حروف کے اعداد کا مجموعہ بتائیں؟	14	16	18	20
.27	اگلا عدد بتائیں 15, 75, 375	5	4	3	6
.28	اگر 2 = 2 , 6 = 3 , 12 = 4 تو 5 = ...	16	18	20	22
.29	اگر 5 برابر ہو 2 کے تو غ کس کے برابر ہوگا؟	ع	ظ	ف	ط
.30	خالی حصہ مکمل کریں (25, 20, 5) (6 , ... , 40)	33	35	34	36
.31	اگلا عدد بتائیں 8 ب 6 ت 4 ث	1	2	3	4
.32	15 کس عدد کا 75% ہوگا؟	20	5	15	75
.33	اگر باپ کو حروفِ تہجی کی ترتیب میں لکھیں تو درمیانی حرف کونسا ہوگا؟	ب	ا	پ	ت
.34	کس عدد کو 12 سے نکالا جائے تو 21 کا تیسرا حصہ حاصل ہو؟	3	5	7	9
.35	حل بتائیں $\sqrt{.09}$ = ...	.03	.9	.3	.09
.36	اسلم کے پاس نوید سے ایک تہائی اور احمد سے دوگنی رقم ہے۔ اگر نوید کے پاس بارہ روپے ہوں تو احمد کے پاس کتنے روپے ہوں گے؟	8	6	4	2

جوابی فارم

رول نمبر : _____ نام : _____ تاریخ پیدائش : _____ جماعت : _____ حاصل شدہ نمبر : _____

کالج : _____ میٹرک تک تعلیم : اردو / انگلش : _____ زیادہ تر رہائش : شہری / دیہی : _____

حصہ اوّل - الفاظ کے معانی

وقت = 12 منٹ

س	ج	س	ج	س	ج	س	ج	س	ج	س	ج
1		8		15		22		29		36	
2		9		16		23		30		37	
3		10		17		24		31		38	
4		11		18		25		32		39	
5		12		19		26		33		40	
6		13		20		27		34		41	
7		14		21		28		35		42	

حصہ دوئم - زبانی اِستدلال

وقت = 6 منٹ

س	ج	س	ج	س	ج	س	ج	س	ج		
1		5		9		13		17			
2		6		10		14		18			
3		7		11		15		19			
4		8		12		16		20			

حصہ سوئم - عددی اِستدلال

وقت = 16 منٹ

س	ج	س	ج	س	ج	س	ج	س	ج	س	ج
1		7		13		19		25		31	
2		8		14		20		26		32	
3		9		15		21		27		33	
4		10		16		22		28		34	
5		11		17		23		29		35	
6		12		18		24		30		36	

حصہ چہارم - معلومات

وقت = 6 منٹ

س	ج	س	ج	س	ج	س	ج	س	ج	س	ج
1		6		11		16		21		26	
2		7		12		17		22		27	
3		8		13		18		23		28	
4		9		14		19		24		29	
5		10		15		20		25		30	