DEVELOPMENT AND STANDARDIZATION OF AN INDIGENOUS NONVERBAL TEST OF INTELLIGENCE

BY

ASHFAQ HUSSAIN GARDEZI

A dissertation submitted to the

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

In partial fulfillment of the requirements for the

DEGREE OF PHILOSOPHY

IN

PSYCHOLOGY

2001

CERTIFICATE

Certified that Ph.D Dissertation titled **"DEVELOPMENT AND STANDARDIZATION OF AN INDIGENOUS NONVERBAL TEST OF INTELLIGENCE"** prepared by **Ashfaq Hussain Gardezi** has been approved for submission to Quaid-i-Azam University, Islamabad.

Prof. Mah Nazir Riaz
supervisor

DEVELOPMENT AND STANDARDIZATION OF AN INDIGENOUS NONVERBAL TEST OF INTELLIGENCE

BY

ASHFAQ HUSSAIN GARDEZI

Approved by

Supervisor

Director, NIP

External Examiner

Dedicated to:

My parents & teachers whom I owe

all my success in life

CONTENTS

ACKNOWLEDGEMENTS

Thanks to Almighty Allah, most benevolent and ever merciful, who bestowed me with the strength and patience to complete this research work.

I am greatly indebted to my supervisor Professor Mah Nazir Riaz, for her valuable advice, expert supervision and constructive criticism, which helped me to complete this task. Without her keen interest and encouragement, it would have not been possible to complete the work in time.

I am grateful to Director NIP, Professor Naeem Tariq, Professor Muhammad Pervez, and other faculty members of the Institute for their cooperation, guidance and encouragement.

I extend my deep gratitude to Dr. Iffat Dar, Dr. Arif, Dr. Tanvir and Syed Sajjad for their useful suggestions and inspirations and help in data collection. I would cordially thank to all my colleagues and staff members of PA-4, GHQ for their sincere cooperation at every step of my research work.

I am also thankful to Mr. Usman for his help in data analysis, Mr. Abdul Qayoom for compilation of thesis, Mr. Hanif and other admin staff for cooperation and provision of necessary support in the completion of this research work.

At the end, I owe my deep appreciation to my wife, Raana Gardezi, son, Ali and daughter, Aamena for their constant support, motivation and prayers throughout my research work. I hardly find suitable words to express my feelings and love for them.

Ashfaq Gardezi
September, 2001

ABSTRACT

The study was aimed at developing and standardizing an indigenous nonverbal test of intelligence for the assessment of general intelligence of Pakistani youth. The items of the test were developed using figurative material taken from our own cultural and folk heritage and were constructed in the framework of English script writing (from left to right). The test comprised of five subtests viz., Series, Matrices, Analogies, Odd One Out and Similarities. It was planned in a manner so as to measure various aspects of intelligence including abstract reasoning, spatial relations, conceptual ability, accuracy of discrimination, eduction of relations and correlates. The effectiveness of the items was judged through item analysis by administering the test to a sample of 200 subjects of grade 12. In evaluating the items, three major aspects; internal consistency, discrimination index and level of difficulty were considered. Only those items were retained that were internally consistent ($p < .001$). Items having discrimination index less than .3 and level of difficulty below .3 and above .7 were discarded. The final draft of the test comprised 90 items; the first three subtests consist of 20 items each and the last two 15 items each. In the main study, the test was administered to 200 subjects; 100 urban and 100 rural. Three methods were used to establish reliability of the test viz., KR-20, split half and test-retest. The estimated indices of reliability were, .89, .85 and .90 respectively. To determine validity of indigenous nonverbal test of intelligence, different validity criteria were used including grade/age differentiation, correlation of the test with other measures of general ability (construct validity) and marks in HSSC examination (criterion-related validity). Significant differences in the mean scores of three grade/age groups on the subtests and the full test, except similarities subtest, were found. To determine the construct validity of the test, factor analysis and the convergent and discriminant validation approaches were adopted. To find out convergent validity, the test was correlated with Intelligence Test Battery (ITB) and adapted version of Raven's Standard Progressive Matrices (RSPM). The results demonstrate significant correlations between all these tests ($p < .001$). The discriminant validity was determined by correlating the test with Individual Obsticles Test, which was insignificant. Criterion-related validity of the test is also high (.74; $p < .001$). No significant differences emerged between urban and rural groups. Similarly, medium of *instruction (English/Urdu) does not seem to influence the test scores. Nevertheless, significant differences were found between Science and Arts groups. Three types of norms viz., Percentiles, T scores and the Deviation IQ were developed by administering the test on a sample of 1000 subjects representative of Pakistani youth.*

LIST OF TABLES

LIST OF ANNEXURES

INTRODUCTION

INTRODUCTION

The measurement of intelligence represents one of the most widespread applications of psychology in every day life (Wigdor & Garner, 1982). Intelligence or IQ testing has been an area of immense interest of psychologists for the last nearly one hundred years. Intelligence tests range from those that tap the general mental ability or general intelligence to those that tap specific abilities such as spatial visualization. Intelligence is measured in a wide variety of applied contexts, often where the results and decisions are of prime significance for individuals, as in selection for jobs or educational courses (Kline, 1991).

Prior to World War II, schools and colleges were the largest users of intelligence tests. During and after World War II, numerous intelligence tests were developed and administered for personnel selection in all branches of military services. The Armed Forces, Public Service Commissions and educational institutions must now be regarded as the major users of intelligence tests.

Historical Perspective

The attempt to measure human intelligence has entailed continuous, long term, intensive efforts than any other project in psychological measurement. In the study of intelligence, philosophical curiosity interacts with practical demands. For centuries men have remained puzzled over the enormous differences in sheer intellectual capacity that separate Socrates from an ordinary citizen, an idiot from a normal child. During twentieth century, psychologists demonstrated a strong concern to investigate if such differences are innate or acquired, whether biological or educational factors are more influential in producing them. Eventually, because of the practical needs of civilized society, it became more and more imperative to find some way to evaluate the intelligence of individuals as accurately as possible. The arrival of universal education brought all sort of children, talented, average and even mentally retarded into the schools. Some appeared to be incapable of mastering the curriculum that had been planned for them. Others raced through it and busied themselves with scientific experiments and philosophical speculations long before they reached physical maturity. Clearly, teacher needed to be able to distinguish between different mental capacities in order to educate children accordingly. Similar problems arose in military organizations and industries regarding personnel selection.

The fact that people differ in abilities, personality, and behavior and that these differences can be assessed in some way has probably been recognized since the dawn of recorded history. Plato and Aristotle wrote about individual differences, but the Mandarins of ancient China preempted even them. As early as 2200 B.C., a civil-service system was instituted by the Chinese emperor to determine if his officials were fit to perform their governmental duties. Later Chinese emperors continued this system, according to which officials were examined every three years. Many centuries later, British and French officials in the 1900s patterned their civil-service examination procedures after the ancient Chinese system (DuBois, 1970; Thorndike, 1990).

Interest in individual differences, at least from a scientific point of view, was almost nonexistent in Europe during the Middle Ages. In the social structure of medieval European society, the class into which s/he was born, providing little freedom for personal expression or development, dictated a person's activities. By the sixteenth century, however, European society had become more capitalistic and less doctrinaire; the idea was growing that people are unique and are responsible for asserting their natural gifts and improving their situations. Thus, the Renaissance can be viewed not only as a period during which interest in learning was reawakened but also as a rebirth of individualism. The spirit of individualism, which flourished with the political and economic stimulation provided by capitalism and democracy, found expression in art, science, and government. It was not until the nineteenth century, however, that the scientific study of individual differences actually got underway.

Early in the nineteenth century, scientists generally viewed individual differences in sensorimotor and mental abilities as more of a nuisance than anything else. Before the prevention of precise, automatic equipment for measuring and recording physical events, the accuracy of scientific measurements of time, distance, and other physical variables depended to a large extent on the perceptual abilities of human observers, who were usually highly trained and very careful in making such measurements. However, despite such precautions the measurements made by different people and by the same person on different occasions were varied. Since the search for general laws of nature is difficult when measurements are unreliable, physical scientists directed their attention to the construction of instruments that would be more consistent and precise than unaided human observations.

Stimulated by the writings of Darwin on the origin of species and by the emergence of scientific psychology, interest in the study of individual differences grew during the latter half of the nineteenth century. Darwin was an English man, but scientific psychology was actually inaugurated in Germany during the last quarter of the nineteenth century. It was during that time that Ebbinghaus, Wundt and other German experimental psychologists demonstrated that psychological phenomena could be expressed in quantitative and rational terms. Occurrences in France and United States were also important to psychological testing. The research of French psychiatrists and psychologists on mental disorders influenced the development of clinical assessment techniques and tests, and the increased attention given to written examinations in American schools resulted in the development of standardized achievement tests and scales. As is true of the history of any field, many people in several countries played significant roles in the pioneer phase of mental measurement. Especially important during the late 1800s were Galton, J. M. Cattell, and Binet. Galton became interested in the hereditary basis of intelligence and in techniques for measuring abilities. A particular concern of his was the inheritance of genius. He also devised a number of simple sensorimotor tests to measure individual differences like tests of reaction time, discrimination, insight, memory, judgement of length, and so on. Using his simple tests, Galton collected measurements on over 9,000 people ranging in age from 5 to 80. Among his many methodological contributions was the technique of correlations, which has continued to be a popular procedure for analyzing test scores.

J. M. Cattell, was an American who, on returning from Germany after taking his Ph.D. in psychology under Wundt, stopped over in England and became acquainted with Galton's methods and tests while serving as his assistant. Later, at the University of Pennsyllvania, Cattell tried relating scores on these simple mental tests of reaction time and sensory discrimination to school marks. The correlations, however, were very low, and it remained for the Frenchman Binet to construct the first mental test that proved to be an effective predictor of scholastic achievement.

The psychologist, Binet and his physician-associate, Theodore Simon were given task in 1904 to develop a method for identifying children who could not profit sufficiently from instructions in regular school classes. For this purpose Binet and Simon constructed an individually administered test consisting of 30 problems arranged in order of ascending difficulty. The problems on this first workable intelligence test, which was published in 1905, emphasized the ability to judge, understand, and reason. A revision of the test, containing a large number of subtests grouped at age levels from 3 to 13 years, was published in 1908. It was the 1908 revision of the Binet-Simon Intelligence Scale that introduced the concept of mental age as a way of expressing an examinee's score on the test. A further revision of the Binet-Simon scale, published in 1911 after Binet's death, extended the test to the adult level. Several revisions followed, but it was the 1916 revision by Terman in the form we still know as the Stanford-Binet Test.

There were many other pioneers in mental testing: Spearman, Wechsler, Cattell and Thorndike, to name a few. The work of Arthur Otis on paper-and-pencil intelligence tests led directly to the construction of two group tests that came to be known as the Army Alpha and the Army Beta devised by a committee of psychologists during World War 1. These two tests, the Army Alpha for literates and the Army Beta for illiterates, were administered to large groups and were used to measure the mental abilities of thousands of American soldiers during and after the war. Hundreds of psychologists and graduate students in psychology were recruited to administer the tests. The application of these group tests of intelligence far outran their technical improvements. Due to indiscriminate use, the tests failed to meet the unwarranted expectations and thereby generated skepticism. Consequently the intelligence tests had been under lot of criticism in American psychology thereafter.

Wechsler developed his tests in response to many of the criticism on the Binet tests. In 1939, he introduced Wechsler-Bellevue Intelligence Scale followed by revision in 1955 known as Wechsler Adult Intelligence Scale (WAIS) and then WAIS-R (1981) being the first reliable test still much in use. Since that time there have been many intelligence tests produced, some specifically aimed at reducing cultural and background effects.

The Concept of Intelligence

The concept of intelligence is as old as the history of mankind. Individual differences in ability that do not stem directly from differences in sensory and motor functions have been noticed by virtually all humans at all times and places. There are variations in the use of nomenclature for intelligence from time to time. The earliest known literature makes use of some concepts of intelligence and recognizes individual differences in intelligence. As Jensen (1980) has pointed out, the concept of intelligence can be found in the great texts of the Hindus and the Ancient Greeks. This is hardly surprising since in almost every activity we can see things being done intelligently or otherwise. Greek Philosophers, Plato and Aristotle drew a distinction between the cognitive aspects of human nature (those concerned with thinking, problem solving. mediating, reasoning, reflecting and so on) and the hormic aspects of human behaviour (those concerned with emotions, feelings, passions and the will). Finally, Cicero, coined the term intelligence, which is still used to refer to a person's cognitive powers and intellectual abilities.

Having created the concept of intelligence, the Greeks went on to make other important contributions. Aristotle contrasted the observed activity or behaviour of a person with some hypothetical underlying capacity or ability on which it depended. The concept of ability is sometimes called a latent structure concept that accounts for the abilities observed. Intelligence seems to be a singular latent structure, which has to be deduced from observed behaviour using the rules of scientific experimental procedure.

The great orator and educator of ancient Rome, Quintillion gave the following advice to teachers, which looks much like something one might read about individual differences in a modern textbook of educational psychology.

"It is generally and not without reason, regarded as an excellent quality in a master to observe accurately differences of ability in those whom he has undertaken to instruct, and to ascertain in what direction the nature of each particularly inclines him; for there is a talent an incredible variety, and the forms of mind are not less varied than those of bodies" (as cited in Stoddord, 1943).

Thomas Aquinas defined intelligence as the power to combine and separate i.e., the ability to see the similarity among dissimilar things and the dissimilarities among similar things, which is a fair characterization of a good many of the items found in present day intelligence tests (as cited in Pyle, 1979).

Throughout the history of psychology, the term intelligence has been used with a wide diversity of meanings, not only by the general public, but also by the members of different disciplines, such as biology, philosophy, education (Sternberg, 1990) and by psychologists who specialize in different areas or identify with different theoretical orientations (e.g., Anastasi, 1983d, 1986c; Brody, 1992; Gardner, 1983, 1993; Glazer, 1993; Lubinski & Benbow, 1995; Messick, 1992; Rowe, 1991; Sternberg & Detterman, 1986).

An early demonstration of this diversity of meanings was provided in 1921, when the editor of the Journal of Educational Psychology invited 17 leading investigators to contribute their definitions and concepts of intelligence. A similar survey was undertaken 65 years later (Sternberg & Detterman, 1986). An examination of these publications is of considerable theoretical interest and provides a basis for lively discussion and possibly some convergence among conflicting views (as cited in Anastasi & Urbina, 1997).

Various thinkers, researchers and psychologists concerning the basis of difference among individuals in the ability to act intelligently have offered various explanations and definitions. Some of these definitions as advanced by some of the renowned psychologists/educators are as follows:

1. Binet (1910) defines intelligence as an ability to judge well, to comprehend well and to reason well.
2. Terman (1916) defines it as the capacity to form concepts and to grasp their significance.
3. According to Spearman (1927), general intelligence involves mainly the eduction of relations and correlates.
4. Stodard (1943) calls it the ability to undertake activities that are characterized by: difficulty, complexity, abstractness, economy, adaptiveness to goal, social value and the emergence of originals, and to maintain such activities under conditions that demand a concentration of energy and a resistance to emotional force.
5. Burt (1954) views intelligence as an innate, general, cognitive ability.
6. Wechsler (1958) considers intelligence as an aggregate or global capacity of the individual to act purposefully, to think rationally and to deal effectively with the environment.
7. Vernon (1960) stresses a simple and non-specific definition, such as 'all round thinking capacity' or 'mental efficiency'.
8. According to Heim (1975), intelligent activity consists in grasping the essentials in a situation and responding appropriately to them.
9. For Piaget (1976), it is adaptation to social and physical environment.
10. Humphreys (1979) considers intelligence as the entire repertoire of acquired skills, knowledge, learning set and generalization tendencies considered intellectual in nature that are available at any one period of time.

11. Cronbach (1984) regards it as a style of work. To say that one person is more intelligent than another means that he acts more intelligently, most of the time.
12. Sternberg (1985) 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.
13. According to Huffman, M. Vernoy, & J. Vernoy (1994), intelligence refers to the cognitive abilities employed in acquiring, remembering, and using knowledge of one's culture to solve everyday problems and to readily adapt to and function in both changing and stable environments.

While comparing these definitions, one finds various theorists emphasize different aspects of intelligence. Some stress the ability to adapt to the demands of the environment; the others believe it is one's ability to learn; some psychologists emphasize the measurement of the ability to reason and other cognitive functions, still others the development of those functions, and probably the layman would mumble something about 'commonsense'.

Nature of Intelligence

The following main approaches have been used by the researchers to study the nature of intelligence.

1. Differential or Psychometric approach
2. Developmental or Piagetian approach
3. Cognitive or Information processing approach
4. Neurological approach

Differential or Psychometric Approach

Differential or psychometric approach to the study of intelligence begins with the assumption that nature of intelligence can best be investigated by studying the ways in which people differ in performance on the tests of intellectual abilities. There is no doubt that people vary widely in many ways, such as their abilities to learn and use words, to solve arithmetic problems, and to perceive and remember spatial information. But the question remains whether intelligence is a global trait or a composite of separate, independent abilities. The fact that psychometricians have devised tests that yield single intelligence or IQ score does not in itself mean that intelligence is a single general characteristic. Some investigators have suggested that certain intellectual abilities are completely independent of one another. For example, a person can be excellent at spatial reasoning but poor at solving verbal analogies. Even investigators who believe that intelligence is a global trait acknowledge that people also have specific intellectual abilities and that these abilities are at least somewhat independent. However, there is still disagreement between those who believe that the specific abilities are totally independent and those who believe that one general factor influence them.

Spearman (1927) proposed two factors model of intelligence, which he felt could explain the pattern of correlation among the groups of intellectual or cognitive tests, he analyzed. According to this model, a person's performance on a test of intellectual ability or on a cognitive task is determined by two factors: the general factor (g) and specific factor (s), which is a factor that is specific to particular test and task. In psychometry, general intelligence is equated with this general factor, (g), and is common to all intellectual tests, cognitive tasks and other mental activities.

Spearman has not been alone in his belief in the explanatory power of a general factor. Binet and Terman demonstrated its existence in their work. There is evidence that performance on the Stanford-Binet and similar tests can be explained largely in terms of "g" factor (McNemar, 1942).

Recent research has also supported two factors model. In a study, Dai et al., (1990) supported general intelligence factor (g) and the Full Scale IQ. They conducted study on a sample of diagnosed brain damaged individuals. Applying factor analysis technique to study the structure of intelligence, they identified two main factors: Verbal-Comprehension and Perceptual-Organization. In another study, the results supported two factors approach as more informative and viable (Piedmont et al., 1992). Enns and Reddon (1998) have also reported that there is a strong evidence for the general factor (i.e., g) and two factors (i.e., verbal and perceptual organization) solutions for the structure of intelligence.

In contrast to Spearman, Thorndike (1924) presented a multifactor model of intelligence. According to Thorndike there is really no such factor as "general intelligence". Rather, there are many highly specific acts, the number of such depend upon how refined classification one might wish to make and is capable of making. Thorndike's is really an "atomistic" theory of mental ability. He adds, however, that certain mental activities have so many of their elements in common that it is useful to classify these tasks into separate groups to which special names are given e.g., verbal meaning, arithmetical reasoning, comprehension, visual perception of relationships, and others. Thorndike has also devised a test to measure ability to deal with abstractions. His test known as CAVD test, is composed of four parts: Sentence completion (C), arithmetical reasoning (A), Vocabulary (V), and following direction (D).

Thurstone (1938), one of the most famous names in psychological and educational measurement, made many methodological and substantive contributions to the field. He proposed a group factor model of intelligence. Applying the method of factor analysis, he extracted seven important group factors that he called as primary abilities and labeled them as: V(verbal reasoning), N(number facility), R(inductive reasoning), P(perceptual speed), S(spatial relations), M(memory) and W(verbal fluency). Most cognitive tests represent complex combinations of these factors. Thurstone and his associates constructed a series of tests, the Primary Mental Abilities Tests, to serve as relatively pure measure of each factor.

Guilford (1988) in his Structure of Intellect (SI) model, proposed that there are 180 independent abilities, each characterized by an intersection of one of six mental operations (cognition, memory recording, memory retention, divergent thinking, convergent thinking and evaluation) on one of five contents (visual, auditory, symbolic, semantic and behavioral) to produce one of six products (units, classes, relations, systems, transportation and implications). This is a three dimensional model with 6x5x6 intersecting cells, each representing an independent ability, hence the 180 factors. Later versions of the theory proposed even more types of intelligence. Although Guilford and Hoefner (1971) claim to have identified 98 of the 180 factors suggested by the model (Guilford, 1988), the theoretical and practical implications of Guilford's model are somewhat in doubt (Allinger, 1988; Bachelor, 1989; Ulosevich, Michael, & Bachelor, 1991). To date, Guilford's model has not greatly affected the intelligence testing.

Vernon (1965) in his hierarchical model of intelligence presented a general cognitive factor (g) at the highest level with two major group factors: the verbal-educational and practical-mechanical-spatial. These major group factors have been further broken down into a number of minor group factors like verbal fluency, numerical ability, mechanical knowledge, psychomotor ability and spatial ability. The lowest levels of the hierarchy are specific factors peculiar to certain tests. Vernon retains the general intelligence factor (g) of Spearman while relegating Thurstone's primary mental abilities and Guilford's structure of intellect factors to a subordinate status under "g". Integrated models of the sort represented by Vernon's hierarchy offer a plausible way of combining the various findings and interpretations of factor analytic research into a single theory.

In a recent study, Arnau and Thompson (2000) have strongly supported Vernon's model of intelligence. Using the standardization sample for the WAIS-III (N = 2450), they observed supporting evidences for the hierarchical factor structure with a second-order factor of intelligence and four first-order factors of Verbal Comprehension, Perceptual Organization, Working Memory and Processing Speed.

Cattell (1971, 1987) performed second order factor analysis and found that general intelligence (g) is composed of two factors; fluid intelligence (g_f) and crystallized intelligence (g_c). Fluid intelligence is defined by relatively culture fair tasks, whereas crystallized intelligence is defined by tasks that require people to have learned information from their culture; particularly, vocabulary and kinds of information that are learned in schools. Cattlell regards fluid ability as being closely related to a person's native capacity for intellectual performance. In other words, it represents a potential ability to learn and solve problems. In contrast, he regards crystallized ability as what a person has accomplished through the use of his or her fluid intelligence, what s/he has learned is based on social, cultural and formal school learning. Horn (1978) differs with Cattell; he cites evidence suggesting that both factors are learned but are also based on heredity. He says that g_f is based on casual learning, while g_c is based on cultural and school type learning.

Fluid intelligence reflects the level of intellectual competence associated with casual learning processes and is assessed by performance on novel, usually nonverbal tests. Crystallized intelligence, on the other hand, reflects intellectual competence associated with international learning processes and is assessed by measure of acculturated knowledge and skills such as verbal tests. (Cattell, 1971; Horn, 1981).

Cattell found that a person's fluid intelligence as measured by culture fair tests reaches its peak in late adolescence or in early adulthood, while crystallized intelligence goes on developing throughout adulthood (Horn & Cattell, 1966, 1967; Lerner, 1990; Willis & Nesselroade, 1990).

Testing both fluid and crystallized abilities shows that the two are most highly correlated in infancy. As children grow up and influenced by their experiences at home and school, so the two abilities diverge (Kline, 1991).

Gustafsson (1984) proposed second-order factor model in which five first-order factors yielded a single second-order factor of general intelligence (g). It claims that fluid ability is essentially Spearman's g. In addition to fluid and crystallized intelligence, perceptual speed, general fluency and broad visualization are also important.

Many research studies now support a hierarchical factor structure of intelligence that is consistent with Gustafsson's second-order factor model (Carrol, 1993; Cattell & Horn, 1978; Gustafsson, 1984; Undheim, 1976, 1978,1981; Undheim & Gustafsson, 1987). In the second-order factor model, five first-order factors including Fluid Intelligence, Crystallized Intelligence, General Fluency, Perceptual Speed and Broad Visualization yield a single second-order factor of Geneal Intelligence (g). The hierarchical nature of second-order factor model is consistent with and reconciles the competing theories proposed by Spearman, Thurstone and Cattell.

Developmental or Piagetian Approach

The developmental or Piagetian approach to the concept of intelligence studies the ways in which infants learn to perceive, manipulate, and think about the world. The most influential proponent of this approach was Jean Piaget.

Piaget's (1952) theory of intellectual development stands markedly in contrast to the psychometric and differential approaches, which are concerned with how people differ with the intent to discover underlying factors or to quantify intellectual skills and identify individual differences. Piaget's approach is more concerned with qualitative aspects of intelligence and with establishing universal pattern such as invariant order of acquisition.

One of the most significant contributions of Piaget in the development of thinking processes of children is the demonstration that the intellect of the child is fundamentally different from that of adult (Flavell, 1993). He explained that a baby begins at a cognitively "primitive" level and that all subsequent intellectual growth progresses in distinct stages motivated by an innate need to know.

For Piaget, the nature of intelligence is a process of organization and adaptation which is forever changing and is not an entity or quantity. He sees intelligence as the way individual adapts to the environment and although emphasis is given to experience in contributing to development, he argues that this is firmly based biologically in the maturational processes that are programmed in the brain. Thus, Piaget thinks that we all have the same intellectual path to follow and this path takes us through a progressive series of cognitive or mental structures or systems which leads to series of stages of development.

According to Piaget, a child comes to know and understand his/her environment by interacting with, and adapting to it. This process is referred to as equilibrium. Equilibrium involves the process of assimilation and accommodation. Assimilation consists of filling new experiences into pre-existing mental structure or Schema. Accommodation is the modification of these schemas in the light of experience. Young children assimilate whenever they take in or use their environment. The process of accommodation occurs whenever the environment reacts. As the child matures, the grasping schema or other mental structures and the associated patterns become elaborated and refined in response to experience. Assimilation and accommodation take place continuously throughout the life while equilibrium occurs only during large scale transition.

Piaget holds that cognitive growth takes place through assimilation and accommodation of the external world, in four stages. During the first period, *Sensorimotor stage* (birth to 2 years), children explore the world and develop their schemata primarily through their senses and motor activities. One of the most important concepts acquired during this stage is object permanence. In the second period, *Pre-operational stage* (2 to 7 years), children have acquired object permanence and can now understand language and other symbolic representation. It is a concrete, egocentric period. During the *Concrete operational stage* (7 to11 years), children develop organized systems of operations by the process of social interaction, with corresponding reductions in self-centeredness. The final period in Piaget's theory is the *Formal operational stage* (11 to 15 years). In this stage children begin to apply their operations to abstract concepts, in addition to concrete objects. They become also capable of hypothetical thinking, can use logic, verbal reasoning and perform higher-level and more abstract operations.

According to Piaget, the entire sequence of four periods is completed in late adolescence i.e., by the age 16 to 17 years. Intelligence, the ability to solve new problems supposedly declines slowly after age of 15 years. Thus, the terminal age of intellectual development in Piaget's system is identical to that of Cattell's fluid intelligence (Aiken, 1997).

Although Piaget never described post-adolescence development of cognitive behaviour in much detail, several neo-Piagetians have suggested that some adults may enter a fifth or later cognitive stage (Arlin,1984; Commons et al., 1989). These researchers suggest that thinking during the formal operation period is primarily a passive intellectual exercise, whereas, thinking in later years is active and employs logic to tackle problems in the real world.

Although Piaget's model is thought to be internally consistent and based on logical and mathematical lines and has dominated twentieth century child psychology, yet not all of his conclusions have been uncritically accepted. Perhaps the major criticism is on his stages of development. Research has shown that Piaget may have underestimated young children's cognitive development. During the sensory stage, for example, later research with more sophisticated equipment has shown that infants develop object permanence much earlier than Piaget suggested (Baillargeon, 1991; Spelke, 1988). Similarly research on the possibility of infant imitation of facial expression also question about Piaget's estimate of early infant cognition. Meltzoff and Moore (1989) suggested that newborns can imitate such facial movements as tongue protrusion, mouth opening and lip pursing (Huffman, Vernoy, & Vernoy, 1994)

Studies of egocentrism in preoperational children have also found that preschool children do show an ability to take another perspective when the testing situation is familiar and the research method simplified (Flavell, 1993; Klemchuk, Bond, & Howell, 1990; Sugarman, 1987).

Piaget's model, like other stage theories, has been criticized for its relative inattention to differences in educational and cultural experiences (Berry et al., 1992; Brislin, 1993; Chance & Fischman, 1987; Cole & Cole, 1989; Flavell, 1985). Formal education and specific cultural experiences can significantly affect cognitive development.

Cognitive or Information Processing Approach

The cognitive or information processing approach to the study of intelligence represents an expansion of cognitive psychology into a field formerly dominated by psychometricians. Psychometric approach deals with intelligence primarily in its structural aspects while cognitive approach deals with it primarily in its processing aspects. The fundamental unit of analysis in most psychometric approaches is factor, while the fundamental unit of analysis in most cognitive approaches is the information-processing component. A component is an elementary information process that operates on internal representation of objects or symbols (Newell & Simon, 1972; Sternberg, 1977). The component is a unit of process just as the factor is a unit of structure. Cognitive psychologists are more interested in the things people do when they solve the tasks.

According to an information processing perspective, to understand intelligence, one must understand the cognitive strategies used by individuals who score high or low on this dimension. In other words, intelligence must be defined in terms of basic aspects of cognition (Matarazzo, 1992; Naglieri & Das, 1990).

According to Carroll (1980) performance on mental tests can be understood in terms of relatively small number of basic underlying information processing components.

Carroll has investigated the major tests used in both psychometric and cognitive research. Based on a logical and partly intuitive analysis of the task, Carroll (1981) has identified a tentative list of ten types of cognitive components. They are monitoring, attention, apprehension, perceptual integration, encoding, comparison, co-representation formation, co-representation retrieval, transformation, and response execution. Although Carroll is not certain that these processes are all mutually distinct from one another, they seem to be different enough to serve as the basis for an information-processing analysis of intelligent task performance.

Brown (1978) has divided process of cognition into two kinds: meta cognitive processes, which are executive skills used to control one's information processing, and cognitive processes which are non executive skills used to implement task strategies. In Brown's particular version of this process dichotomy, five meta cognitive processes are of particular importance. Planning one's next move in executing the strategy, monitoring the effectiveness of individual steps in a strategy, testing one's strategy as one performs it, revising one's strategy as the need arises and evaluating one's strategy in order to determine its effectiveness.

Sternberg (1980) distinguishes three different kinds of information processing components: Meta components, performance components and knowledge-acquisition components. Meta components are higher order control processes used for executive planning, monitoring and evaluation of one's performance in a task. Meta components are believed to be most important in intelligent functioning. Performance components are lower-order processes used in the execution of various strategies for task performance. Knowledge-acquisition components are processes involved in learning new information and storing it in memory. The three knowledge-acquisition components believed to be most important in intelligent functioning are: selective encoding, selective combination and selective comparison. These three kinds of components are applied in task performance for reaching a solution or other goals.

Sternberg (1980) has described four ways in which various kinds of components can interact with each other. First, direct activation of one kind of component by another. Second, indirect activation of one kind of component by another via the mediation of third kind of component. Third, direct feedback from one kind of component to another. Fourth, indirect feedback from one kind of component to another. In the proposed system only meta components can directly activate and receive feedback from each other. Thus all control passes directly from the meta components to the system, and all information passes directly from the system to the components. The other kinds of components can activate each other indirectly, and receive feedback from each other indirectly; in every case mediation must be supplied by the meta components.

Sternberg's notion of intelligence is extremely wide. It also concerns with the aspects of motivation and includes social competence. Nevertheless, Sternberg's work is effective in explaining the nature of intelligence as defined by the tests of intelligence. He used analogical reasoning in his componential theory because it is one of the best measures of 'g' factor. In his model of analogies, five components: encoding, inference, mapping, application and response were identified. Encoding involves translation of stimulus into a mental representation. Inference is to find the rule according to which the first two figures or patterns (problem part) of the item are related to each other while mapping finds the rule that relates the first two figures (problem part) and the third and fourth figures (answer part) of the item. Application is applying the inferred relations of the first two figures (problem part) to the third and fourth figures (answer part) of the item to arrive at the solution. Sternberg (1986) claims that higher scorers on standard tests of intelligence are faster at inferring, mapping and applying relation, as well as responding which is a component of different type. Pellegrino (1985) in his study of inductive reasoning shows that components are applied in classification and in series completion items of intelligence tests.

Sternberg (1985) basing on findings from information processing research developed what he calls a triarchic theory of human intelligence. He feels that even more important than the outward products of intelligence such as correct answers on an intelligence test or solution to problems are the thinking processes used to arrive at answers to problems, and theories of intelligence should account for these processes. According to the triarchic theory, there are three aspects of intelligence that are separate but related: the internal components of intelligence, the use of these components to adapt to environmental changes and the application of past experience to real life situations (Frensch & Sternberg, 1990). Some people tend to have a stronger aptitude for using one or more of these aspects of intelligence. Sternberg suggests that intelligence depends on acquiring information processing skills and strategies to solve problems and it can not be understood outside a socio-cultural context. What may be relevant in one culture may not be in the other culture.

Sternberg viewed that there are actually three basic types of human intelligence. The first, known as *componential or analytical intelligence* involves the abilities to think critically and analytically. Persons high on this dimension usually excel on standard tests of academic potential. The second type of intelligence known as *experiential or creative intelligence*, emphasizes insight and ability to formulate new ideas. This is the kind of intelligence shown by many scientists, geniuses and inventors. The third type of intelligence is termed as *contextual or practical intelligence*. Persons high on this dimension are intelligent in a practical, adaptive sense. Sternberg viewed that intelligence might be measured by taking someone who is well adaptive to one culture and placing him/her in an unfamiliar one, in order to assess his/her ability to cope with a new situation. The more a person is able to automatize the tasks of daily life, the more mental resources there are left to cope with novelty (Sterberg, 1985; Sternberg et al., 1995).

Neurological Approach

Neurologists have traced the individual differences in intelligence on the basis of differences in neural functioning. Highly intelligent persons are often described as "fast thinkers" in responding to changing situations and new events. This everyday usage points to another possible perspective on intelligence, the one emphasizing neural factors, such as more rapid or efficient processing of information by nerve cells within the brain. This approach has gained increasing attention among psychologists in the recent years. A growing body of evidence suggests that intelligence may actually be closely linked to physiological processes, especially ones going on in the nervous system and in the brain in particular (e.g., Matarazzo, 1992; Vernon, 1993).

Reed & Jensen (as cited in Baron, 1999) recorded evoked potentials, or electrical responses, in the brains of 147 male volunteers who were presented with a visual stimulus. The average latency, or delay, with which these potentials followed presentation of the visual stimuli was obtained for each volunteer; then the latency was divided by the length of the volunteer's head to obtain a measure of the speed with which nerve impulses were conducted in the visual system. The data was then correlated with the volunteer's scores on a test of intelligence, the Raven Progressive Matrices. The results revealed that the faster the neural speed of the participants, the higher was their IQ scores.

Haier (1996) has examined metabolic activity in the brain during cognitive tasks. It was found that more the intelligent people are, the less energy their brains expend while working on various tasks.

Andreasen et al., (1993) demonstrated that there is a link between brain structure and intelligence. More specifically, scores on standard measures of intelligence are related to the size of certain portion of brain, including the left and right temporal lobes and the left and right hippocampus.

From the above studies, it appears that the improved methods now available for studying the brain and nervous system are beginning to establish the kind of links between intelligence and physical structures that psychologists have long suspected to exist. However, the present research is not yet enough to warrant firm conclusions. Still, it does appear that we are on the verge of establishing much firmer links (Baron, 1999).

Multiple Intelligences

Gardner (1983) proposed a theory of multiple intelligences which suggests that there are a number of distinct forms of intelligence that each individual possesses in varying degrees. He defines multiple intelligence as "the capacity to solve problems and to fashion products in a context- rich and naturalistic setting" (Armstrong, 1994, p.2). Gardner proposes seven primary forms; linguistic, musical, logical-mathematical, spatial, body-kinesthetic, intrapersonal (e.g., insight, metacognition) and interpersonal (e.g., social skills). He feels that each person tends to excel in some areas more than the others.

According to Gardner, the implication of the theory is that learning/teaching should focus on the particular intelligence of each person. For example, if individual has strong spatial or musical intelligences, s/he should be encouraged to develop these abilities.

Gardner points out that the different intelligences represent not only different content domains but also learning modalities. A further implication of the theory is that assessment of abilities should measure all forms of intelligence, not just linguistic and logical-mathematical. Consequently, he maintains, intelligence testing should consists of assessing a person's strengths rather than coming up with a single "IQ score" (Gardner, 1986).

Gardner also emphasizes the cultural context of multiple intelligences. Each culture tends to emphasize particular intelligence. For example, Gardner (1983) discusses the high spatial abilities of the Puluwat people of the Caroline Islands, who use these skills to navigate their canoes in the ocean. Gardner also discusses the balance of personal intelligences required in Japanese society.

The theory of multiple intelligences has been focused mostly on child development although it applies to all ages. While there is no direct support for the theory, Gardner (1983) presents evidence from many domains including biology, anthropology and the creative arts.

Gardner (1993) discusses application of the theory to school programs. He maintains that multiple intelligences and learning styles are similar but they begin and end in different places. He states that proponents of learning styles seek to describe an individual in terms of one approach to learning for all the content areas. While the discipline (intelligence) dictates the approach that an individual will use to acquire knowledge. Some educators have taken the seven multiple intelligences and combined them with learning styles. They identify learners as linguistic, logical-mathematical, spatial, body-kinesthetic, musical, interpersonal and intrapersonal (Marks-Tarlow, 1996).

Emotional Intelligence

Salovey and Mayer (1990) first introduced the model of emotional intelligence. Salovey subsumes Gardner's *Personal intelligence* in his basic definition of emotional intelligence, expanding these abilities into five main domains: knowing one's emotions, managing emotions, and motivating oneself, recognizing emotions in others and handling relationships.

Environmental Influences on Intelligence

People differ in intellectual ability. How much of this difference is due to the particular genes one inherit and how much is due to the environment in which one is raised? The heredity-environment issue, debated in regard to many aspects of human behavior, has focused primarily on the area of intelligence. Today, most researchers believe both heredity and environment play a large part in determining people's intelligence (Bouchard et al., 1990; Plomin, 1989; Thompson, Determan, & Plomin, 1991). But for years the debate went back and forth as different studies supported one or the other. The experts agree that at least some aspects of intelligence are inherited, but again opinions differ as to the relative contributions of heredity and environment. However, one can think of a person's genes as imposing a top and bottom limit on intelligence i.e., establishing a range of intellectual ability. Environmental influences i.e., what happens to the individual during the course of development, will determine where within that range the person's IQ will fall. In other words, genes do not specify behaviour but establish a range of probable responses to the environment.

The environmental conditions that determine how the individual's intellectual potentials will develop include nutrition, health, economical and emotional climate of the home, quality of stimulation, education, schooling and the type of feedback given for behaviour. Given two children with the same genes, the child with the better prenatal and postnatal nutrition, the more stimulating and emotionally secure home and school environment, and the more appropriate rewards for academic accomplishments will attain the higher intelligence score when tested on an ability or intelligence test. Studies have shown that intellectual differences between children of lower socio-economic status with malnutrition, poor stimulating and emotionally insecure home and school environment and high socio-economic status, with better nutrition, health, more stimulating and secure home, social and educational environment become progressively greater between birth, entrance into school and adulthood (Bayley, 1970). This suggests that environmental conditions accentuate whatever differences are present at birth.

Weinberg (1989) asserts that human intelligence could be viewed as malleable. This plasticity of human behavior is a pervasive quality of human organism. Individual differences in the malleability of intellectual functions are in part due to genetic make up. But the genes do not fix the behavior; rather, they establish a range of possible reactions and a range of possible experiences that environment can provide. The environment can determine whether the full range of genes activity is expressed. Thus, how quickly we learn depends not only on the genetic endowment but also on the nature of the environment. He also stressed that in order to assess the potential of the intellectual level, one has to provide the optimum level of environment in which a particular skill is assessed.

Though probably not as important a factor as the others, still attending school has some affect on the intelligence of children. Some would argue that important aspects of intelligent behavior are learned outside school (mainly before school begins). Vernon (1969) stresses that the sheer amount of schooling, regardless of quality, helps to promote the kind of reasoning measured in intelligence tests. Pre-school education and early school experiences stretch the cognitive skills of children and in turn strengthen their intellectual power. Ceci (1991) later supported Vernon's findings.

Pidgeon (1970) has demonstrated that children's intelligence scores can be affected quite dramatically by subtle factors operating within the school. He argues that the belief held by teachers about the concept and nature of intelligence determines, to a large extent, the level of achievement expected by pupils. The subtle influences operating within a school are able either to help children develop or effectively to slow down their intellectual development.

There is quite a substantial amount of research concerning the impact of home environment, which shows that certain factors operate in rather subtle, yet quite powerful way in influencing the intelligence of children. Some large studies, following up thousands of children over many years, show that parental interest and encouragement can be important factors in determining ability and attainment scores. In certain cases, these factors of interest and encouragement shown by parents towards their children's schoolwork are more powerful than other factors (Douglas, 1969; Douglas, Ross, & Simpson, 1968). There is also a strong association between parental educational level and the ability of the children (Bouchard & Segal, 1985; Davie, Butler, & Goldstein, 1972).

Flynnn (1994) studied the IQ scores for different populations over the past sixty years and discovered that IQ scores have increased from one generation to the next for all of the countries for which data existed. This interesting phenomenon has been called the *Flynn Effect*. It has opened a new debate among the researchers. Research shows that IQ gains have been mixed for different countries, but in general, countries have shown generational increases between 5 and 25 points. Massive shifts in human heredity, changing environmental conditions, urbanization all over the world, expansion in mass media and improvements both in nutrition and in educational opportunities in many countries have been considered as causal factors for this massive increase.

Restricted cultural environment (e.g., insisting the children to work alone and sit still and quietly in the class) is another possibility that minority persons score lowest (Boykin, 1994). Low self esteem of 'caste like minorities' keep them avoiding to exert themselves on any challenging task. They grow up with low expectations, believing that efforts on their part will not result in better outcomes. This leads them to reject academic achievement and other forms of behaviour described by majority group (Ogbu, 1994).

Intelligence and Academic Achievement

There is a huge number of researches to prove that a reasonably close relationship exists between intelligence scores and academic performance, which is hardly surprising since the origin of intelligence testing lay in the effort to select children worthy of education, for example Binet and Simon (1905). However, as both Snow and Yalow (1982) and Jensen (1980) pointed out, this work is relatively easy to summarize simply because the results are clear, remarkably so by the standards of psychological research.

As Jensen argues, the results in support of a substantial correlation between intelligence test scores and academic achievement are so incontrovertible that critics of the tests accept them. However, they interpret them not as an evidence for a link between intelligence and academic success, but rather as a criticism of the tests themselves: that the tests reflect simple socio-economic advantage, or increased educational opportunities. Nevertheless it has been demonstrated empirically that intelligence measured at the age of five predicts better than any other variable, a child's future academic progress and attainment. To quote Jensen (1980) "Children with higher IQs generally acquire more scholastic knowledge more quickly and easily, get better marks, like school better, and stay in school longer" (p.317). However, these claims require further explication. The first point to note is that intelligence and academic success or achievements are not perfectly correlated. It would be surprising, if they are, since it is obvious that variables other than intelligence play important part in one's academic achievements both at school and college. For instance, hardwork, preference for certain subjects, the quality of teaching, health and home environment are just a few examples of variables other than intelligence that affect significantly one's achievements.

One other argument that might be used against the interpretation of these findings as support for the importance of IQ in determining scholastic achievement is that IQ itself is also affected by all these variables, implying that IQ is no more than learning. However, examination of the content of IQ tests, shows that this interpretation will not hold, especially with tests of fluid ability where items are used which are quite unfamiliar to the subjects. There is no learning involved with Raven's Matrices (Kline, 1991).

Consequent upon lot of research in the field, it is fair to conclude that the substantial correlation between IQ scores and academic attainment can not be explained by arguing that intelligence itself is essentially a learning and an attainment. One further closely related point should be dismissed, that is, the content of the tests is highly similar to what has to be learned in school, and this factor alone accounts for the correlation. In the case of certain intelligence tests like WISC and WAIS, there is some truth in this point. However, it should be noted that the information and vocabulary scales are more than counterbalanced by the other scales in the test, which load also on the g factor. Certainly it is preferable to measure intelligence in a less culturally biased form and this is exactly what the fluid ability tests, developed factor analytically, do (Kline, 1991).

Snow and Yalow (1982) argue that intelligence and learning ability can not be simply equated. There is a complex relationship between them as is shown by the fact that estimated true mental age in any year is substantially correlated with gain in mental age in subsequent years (Cronbach & Snow, 1979). In other words, the child with high intelligence is likely to improve yet more relative to his/her less intelligent peers.

Jensen (1980), from the survey of the research on learning and its relation to IQ scores has demonstrated well that some generalizations can be made. Thus learning does correlate with IQ when it requires conscious mental effort and is intentional, compared, for example, with conditioning processes, which go on without awareness.

Jensen (1980) further points out that learning and IQ are correlated more at the initial stages of learning something new than at later stages where gains are the results of practice. This is best exemplified by music. Learning to read music is positively correlated to IQ. However, it is not true for skilled music practitioners, for whom sight reading is automatic.

Nonverbal Tests of Intelligence

Tests designed to measure one's mental capacity or intellectual achievements or level of general cognitive functioning are called intelligence tests in the psychological literature or IQ tests in popular discussion. These tests are frequently employed as primarily screening inventories to be followed by tests of special aptitudes. This practice is essentially prevalent in the testing of normal adolescents, or adults for educational and occupational counselling, personnel selection and similar other purposes. The other use of general intelligence tests is to be found in clinical testing, especially in the identification and classification of the mentally retarded (Anastasi & Urbina, 1997).

Besides many other classifications, intelligence tests can also be classified as verbal and nonverbal intelligence tests. A verbal intelligence test is one, which involves the use of language. It requires the examinee to read the given exercises or solve problems presented orally. A nonverbal test of intelligence on the other hand, does not necessitate the use of language. Instead of words, it rather utilizes figures, pictures, geometric patterns and designs, symbols and the like. Nonverbal intelligence tests are generally comprised items like series, analogies, classification, matrices, faulty pictures, figure generalization or topological conditions, imbedded figures and Gestalt completion etc. All the nonverbal items are loaded on Spearman's general factor "g". They involve functions like abstract reasoning, spatial relations, eduction of relations, perceptual ability, accuracy of discrimination, permutation and alternation of patterns and other types of logical reasoning.

All the nonverbal tests of intelligence though having common aim of measuring a unitary capacity, are not alike. On some tests the items are mixed up, with a single time limit, while on other tests the items are grouped as sets of separately timed subtests. Nonverbal tests of intelligence are usually classified into individual and group tests. Each has its own advantages and limitations. Individual nonverbal intelligence tests, which are administered to one examinee at a time, are somewhat different in focus from group intelligence tests, which can be administered simultaneously to many examinees. Individually administered nonverbal intelligence tests are generally employed for clinical purposes, while group administered non verbal intelligence tests are used for educational and occupational counselling, personnel selection and other similar purposes.

Individually Administered Intelligence Tests

The Binet-Simon Scale (1905) was an individual test consisting of 30 problems/tests arranged in ascending order of difficulty. The tests covered many areas with special emphasis on judgment, comprehension and reasoning. This scale was later revised several times. Binet test, in the revised forms is still used all over the world for measuring intelligence of infant, children and adults. The most important revisions of Binet scale are: The Stanford Revision and Extension of Binet-Simon Scale by Terman in 1916 and 1937, Revision and Standardization by Terman and Merrill in 1960 and 1973 and Kuhlmann-Binet revision which extended the scale downward to the age of 3 months. The latest, fourth edition of Stanford Binet Intelligence Scale is released in 1986 (Anastasi & Urbina, 1997).

The Binet scale and its revisions are largely verbal in contents, although some nonverbal items are included at early age levels.

The Wechsler scales (Wechsler, 1939) can be considered as an important step in the history of test development that combine both verbal and nonverbal material within a single instrument to obtain the advantages, comparisons and contrasts provided by both types of test items. Unlike the Standford-Binet, there are separate Wechsler scales for adult and children. The Wechsler Adult Intelligence Scales (WAIS) first published in 1955, was revised in 1981 (The WAIS-R). The Wechsler Intelligence Scale for Children (WISC), first developed in 1949 and revised in 1974 (The WISC-R) resembles WAIS and is used for children from 5 through 15 years of age. The Third Edition of WISC (WISC-III) published in 1991, intended for children aged 6 years to 16 years and 11 months. Wechsler also devised an intelligence test for pre-school children, which is known as

Wechsler Pre-School and Primary Scale of Intelligence (WPPSI), published in 1968 and revised in 1989. The WAIS-R consists of 11 subtests, divided into two categories: verbal and performance. The Wechsler tests are currently among the most frequently used individual tests of intelligence. Recently WAIS-III has been developed (Murphy & Davidshofer, 1998).

Pictorial Test of Intelligence (French, 1960) is a nonverbal test that consists of a set of 11 x 11 inch stimulus cards containing four pictures each; the child responds to the examiner's query concerning each card by pointing to or looking at the correct picture. Though an untimed test, it takes about 45 minutes. It is composed of six subtests: *Picture Vocabulary, Form Discrimination, Information and Comprehension, Similarities, Size and Number, Immediate Recall.* Although pictorial intelligence tests are easy to administer and score, they have a serious limitation: because of their multiple-choice format, the score may be higher than justified by an examinee's true mental ability. Nevertheless, these tests of intelligence have been found helpful in evaluating the mental abilities of orthopedically and speech-handicapped children (Aiken, 1997).

Columbia Mental Maturity Scale (Burgemeister, Lorge, & Blum, 1972), has 92 items arranged in eight overlapping levels, but only 51 to 65 items are used in testing a particular examinee. On this scale the examinee selects the drawing that does not belong to any series of pictorial and figure drawings. This test is appropriate for children aged 3 1/2 through 10 years and takes 15 to 20 minutes to administer (Aiken, 1997).

Matrix Analogies Test (MAT) (Naglieri, 1980) is an individually administered nonverbal measure of intellectual ability for ages 5 to 17. It is designed for administering as screening test. It consists of two forms: Short Form (MAT-SF) for use when rapid screening of students is desired and when group administration is possible; Expanded Form (MAT-EF), for use when more in-depth investigation in nonverbal ability is desired.

MAT contains 64 items, and is divided into four 16 items groups: *pattern completion*; *reasoning by analogy*; *serial reasoning* and *spatial visualization.* The items groups were apparently not derived empirically but by *logical organization*. The time limit for the test is 30 minutes (Conoley & Kramer, 1989).

Peabody Picture Vocabulary Test-Revised (Dunn & Dunn, 1981), has 175 pictorial plates arranged in ascending order of difficulty by age level, containing four pictures each. The examiner presents a plate, says a word, and instructs the examinee to point to the picture on the plate that best illustrates the meaning of the word. The test, which takes about 10 to 20 minutes to administer and score, can be employed with a wide range of examinees (2 1/2 years to adulthood). Since no verbal response is required, the PPVT-R can be given to persons who have speech impairments, cerebral palsy, reading problems, and to mentally retarded, withdrawn, or distractible children (Aiken, 1997).

Kaufman Assessment Battery for Children (A. S. Kaufman & N. L. Kaufman, 1983) is another measure of intelligence and achievement which has generated a great enthusiasm among researchers and practitioners of school psychology as the new major individually administered cognitive instrument. The K-ABC was designed primarily to assess the abilities of children aged two and half through twelve and half years to solve problems requiring simultaneous and sequential mental processes. The battery also includes an Achievement Scale to measure acquired skills in reading arithmetic. Based on extensive research on neuro psychology and cognitive psychology, the K-ABC is said to be especially appropriate for determining the mental abilities of pre-school, minority, and exceptional children. Scores in four areas; *Sequential Processing, Simultaneous Processing*, *Mental Processing, Composite* (Sequential plus Simultaneous), and *Achievement* are obtained. Thirteen out of a total of sixteen game-like sub-tests comprising the battery are administered to a given child in a time of 31 to 50 minutes for preschoolers or 50 to 80 minutes for older children (Murphy & Davidshofer, 1998).

Group Administered Intelligence Tests

Group tests were developed at the start of World War I, when the Armed Forces in the United States suddenly faced the problem of screening several million recruits. The first mass testing began during the World War I with the development of the Army Alpha and the Army Beta for use in U.S. Army. The former was verbal test designed for general screening and placement purposes and the later was non language test for use with men who could not properly be tested with the Alpha owing to foreign language background or illiteracy. The pattern established by these tests was closely followed in the subsequent development of a large number of tests for civilian use. Revisions of the civilian forms are still in use as Alpha examination modified form I (Alpha-I) and revised Beta examination (Beta-I).

In U.S. Armed services, the Armed Forces Qualification Test (AFQT) is used as a preliminary screening instrument. AFQT provides a single score based on equal number of vocabulary, arithmetic, spatial relation and mechanical ability items.

The Ohio Penal Classification Tests (Vernon & Scollary, 1957) are group administered, nonverbal, paper and pencil measures of adult mental ability. There are two forms of the OPCT, one for use in prison and the other for use in industry. Each test is composed of 4 subtests. The first three subtests are identical for all forms. Test I includes counting blocks and is supposed to measure spatial perspective attitude and reasoning capacity. Test 2 is an original digit-symbol task purporting to measure perception and associative learning speed. Test 3, requires the completion of 20 number series arranged in order of increasing difficulty and is designed tc measure number facility reasoning. Test 4 consists 10 drawings of common objects whose names the subject is required to write, after a brief exposure, this is supposed to measure both apperception and memory (Buros, 1959).

The Lorge-Thorndike Intelligence Test (Lorge & Thorndike, 1957) is a group intelligence test developed for kindergarten to grade 12. The test contains five scales for five different levels. The two lowest scales for kindergarten to grade 3 contain entirely nonverbal material while three highest scales from grades 4 to 12 are both verbal and nonverbal in nature (Buros, 1965).

The Test of General Ability (Flangan, 1957) is a nonverbal, pictorial measure of intelligence designed to be independent of reading, arithmetic, and scholastic tasks. The test takes about 45 minutes and has a range from kindergarten through grade 12. It is especially devised for children from culturally deprived backgrounds. The score on part I of the test is a measure of the information, vocabulary and concepts that the examinee has acquired. The score on part II is a measure of non- cultural reasoning ability and the total score is an overall, nonverbal measure of general intelligence (Buros, 1965).

The Kuhlmann-Anderson Tests (Kuhlmann & Anderson, 1960) are a modern adaptation of intelligence test devised by Frederick Kuhlmann many years ago. The seven levels of Kuhlmann-Anderson tests extend from Kindergarten through grade 12, each level taking 50 to 75 minutes to administer (Aiken, 1997).

The Figure Reasoning Test (Daniels & Lockwood, 1962) is a nonverbal intelligence test designated for children of 10 and over. It consists of 45 test items (Matrices) and 6 preliminary examples. These test items involve principles like addition, subtractions, deletion, super imposition and progression of various kinds. The time limit for the test is 30 minutes (Buros, 1965).

The Hd Km Nonverbal Test, Forms A and B (Kalyan & Violet, 1962) is designed to measure intelligence which is defined operationally as the ability to interpret and use symbols, the usual and familiar, as well as the unusual and unfamiliar ones. Forms A and B of the test may be given together or independently. Each form consists of 60 items divided into two sections. Part I consists of 23 classifications and Part II 37 analogies (www.ericae.net. on line).

The Goodenough-Harris Drawing Test (Goodenough & Harris, 1963) is a revision of Draw a Man Test, together with a similar Draw a Woman Scale and an experimental Self Drawing Scale. The man and woman figures that the examinee is instructed to draw are scored for bodily and clothing details, proportionality among the various body parts (for example, head-to-trunk) and other characteristics, rather than according to artistic merit. The test is untimed but usually takes 10 to 15 minutes to complete. Although the original hope was that Goodenough-Harris would measure basic intelligence relatively free of cultural influences, it is now realized that the task of drawing a human figure is significantly affected by cultural experiences (Aiken, 1997).

The Otis-Lennon Mental Ability Tests (Otis & Lennon, 1968) are revision of the earliest tests in the Otis series; the Otis Self-administering Tests of Mental Ability and the Otis Quick-scoring Mental Ability Tests. Like their predecessors, the Otis-Lennon tests are composed of a variety of items, both verbal and nonverbal, to measure general ability. The six levels of the tests extend from primary 1 through advanced grades (10-12). Testing time varies from 30 to 45 minutes, depending on the level (Aiken, 1997).

The Herman-Nelson Tests of Mental Ability and its revision (Nelson, Lamke, & French, 1973) cover four grade levels: grade 3-6, 6-9, and 9-12 (Form-1) and K-2 (Primary Battery). A college-level edition of the tests is also available. Each of the three levels of Form I consists of 90 items arranged in spiral-omnibus format, including items on scrambled words, verbal analogies, verbal classification, verbal inference, number series, arithmetic reasoning and figure analogies. Testing time of Form 1 is 30 minutes. The Primary Battery is untimed but usually takes 25 to 30 minutes. It is composed of three subtests: a listening test of thirty general information items, a picture vocabulary test of thirty-five items, and a size and number test of twenty-three items (Aiken, 1997).

The Cognitive Ability Test, Form 3 (Costantino, 1985) is designed for assessing the development of cognitive abilities related to verbal, quantitative and nonverbal reasoning and problem solving (Conoley & Kramer, 1989).

The Schaie Thurstone Adult Mental Ability Test (Schaie, 1985) was developed to measure the mental abilities of adults of age group 22 and above. This test consists of many items of Thurstone Primary Mental Ability Test forms 11-17. This test consists of 7 scales: *Recognition, Vocabulary, Figure Rotation, Letter Series, Number Addition, Word Fluency, Object Rotation and Word Series* (Conoley & Kramer, 1988).

Cognitive Abilities Test devised by Thorndike and Hagen (1987) consists of a Primary Battery for K-3 and a Multilevel Edition (A-H) for grades 3-12. There are two levels of the Primary Battery: Primary 1 (K-2) and Primary 11 (grades 2-3). The Primary Battery and Multilevel Edition are composed of three batteries (Verbal, Quantitative, Nonverbal). Each of the batteries contains two subtests on the Primary Battery Levels and three subtests on the Multilevel Edition. Work time for each subtest is 12-18 minutes on the Primary Battery and 8-12 minutes on the Multilevel Edition. The verbal battery contains vocabulary, sentence completion, verbal classification and verbal analogy sub-tests. The quantitative battery, which includes subtests of quantitative relations, number series and equation building assesses the ability to work with numbers and other quantitative symbols. The nonverbal battery comprised of spatial, geometrical and figural pattern includes subtests of figure classification, figure analogies and figure synthesis (Aiken, 1997).

The Test of Nonverbal Intelligence (Brown, Sherbenou, & Johnsen, 1990) is a measure of reasoning ability. It was developed for persons between the ages of 5 and 85 years and can be administered individually and in groups. The time limit for the test is 15 minutes. The test provides full scale IQ or equivalent and fairness for males and females, African and American, bilingual, hearing impaired and colour-vision deficiency samples (www.ericae.net. on line).

The Comprehensive Test of Nonverbal Intelligence was devised by Hammill and Donald (1996). It is a battery of six subtests that measure different but interrelated nonverbal intellectual abilities. It is designed for persons between the ages of 6 years to 18 years and 11 months. The six subtests are: *Pictorial Analogies, Geometric Analogies, Pictorial Categories, Geometric Categories, Pictorial Sequences and Geometric Sequences*. This test has three principal uses: to assess the intellectual ability of individuals for whom most other mental ability tests are either inappropriate or biased; to make comparisons between verbal and nonverbal intellectual ability; and to use in research studies investigating nonverbal ability and related topics (www. ericae. net. on line).

The General Ability Measure (Naglieri, Bardos, & Achilles, 1998) is a brief self-administered, nonverbal measure of intelligence. The test evaluates an individual's overall general ability with items that require the application of reasoning and logic to solve problems that exclusively use abstract designs and shapes. The instrument contains 66 items comprising four subtests: *Matching, Analogies, Sequences,* and *Construction.* The test can be completed in 25 minutes (www.ericae.net. on line).

The Universal Nonverbal Intelligence Test (Bracken & McCallum, 1998) is a language-free, culture-fair test of general intelligence and specific cognitive abilities. It was specifically developed to assess intelligence of children and adolescents who have language, hearing, or speech impairments or who come from non-English cultures. It is composed of six subtests viz., *Symbolic Memory, Spatial Memory, Object Memory, Cube Design, Analogic Reasoning,* and *Mazes*. It was normed and validated on a large nationally representative sample of children and adolescents living in the United States. An important goal in the development of the Universal Nonverbal Intelligence Test was to create a test that would be reflective of the psychometrically defined "g"(www.ericae.net. on line).

Psychological Issues in Ability Testing

The apparent success of early intelligence tests in classifying recruits during World War 1 led to a boom in civilian intelligence testing in the 1920s and 1930s, and laid the foundation for the large scale intelligence testing industry that exists today. Since its inception, however, large scale intelligence testing has been the focus of sustained public controversy (Cronbach, 1975). Although conflicts between test specialists sometimes revolve around technical issues, the primary issue which has fuelled the bitter public debate over IQ testing has been the fact that there is a difference in the average test scores of different social, racial and ethnic groups of the society (Aiken, 1991; Cohen et al., 1988; Herrnstein & Murray, 1994; Weinberg, 1989). A number of possibilities have been suggested to account for such test score differences. Hereditarians suggest that genetic factors are partially responsible for individual differences in intelligence test scores, while the environmentalists suggest that test scores differences can best be understood in terms of differences in the environments, particularly in the cultures of the examinees. However, most psychologists believe that there is a good deal of evidence to suggest that intelligence is affected by both heredity (DeFries & Plomin, 1978) and environment (Carroll & Maxwell, 1979) and that both factors be taken in to account in understanding individual or group differences in scores on intelligence tests (Murphy & Davidshofer, 1998).

There is a wide spread belief that intelligence tests are biased and that the use of these tests in educational placement, academic admission and personnel selection results in decisions which are fundamentally unfair. One line of evidence which supports this assumption is the existence of systematic differences in test scores as a function of socioeconomic status and sex etc. (Jensen, 1980; Linn, 1982). These differences have been attributed by some authors to test bias and have been cited as an example of the unfair nature of tests (Jackson, 1975; William, 1974). Bias in measurement occurs when the test makes systematic errors in measuring a specific characteristic or attribute. For example, several psychologists claim that IQ tests may be valid for some particular class, but not for other classes (Joseph, 1977; William, 1974). Researchers on test bias have reported that in general:

1. Test scores of children and adults in the middle upper classes tend to be higher than children and adults in the lower socioeconomic classes.
2. Children from high standard educational institutions tend to have higher scores than children from low standard educational institutions.
3. Males as compared to females, receive systematically higher scores on some tests and systematically lower scores on other tests.

Research on differential validity most clearly addresses the question of whether tests measure different attributes for different groups and in that whether it deals with the relationship between test scores and other variables of interest. At one time, the hypothesis that ability tests were differentially valid for predicting criteria such as grades in school, performance on the job, or success in training, was widely accepted (Arvey, 1979; Guion, 1965). The basis for this belief was the fact that many ability tests showed significant correlation with relevant criteria when used with subjects from upper socioeconomic classes, but showed non-significant validity coefficients when used with subjects from lower socioeconomic classes. Although some researchers still maintain that differential validity occurs in specific settings (Katzell & Dyer, 1977; Schmitt, Mellon & Bylenga, 1978), the consensus of most testing specialists is that differential validity is not a widespread phenomenon (Cleary et al., 1975; Hartigan & Wigdor, 1989; Linn, 1982).

Examination of research evidence dealing with group differences in the reliability, the factor structure, or the validity of cognitive ability tests provides little support for the hypothesis that bias in measurement is widespread (Cleary et al., 1975; Jensen, 1980). It is possible however, that these tests are biased or less valid in the predictions they make for lower groups (Aiken, 1991).

In an effort to eliminate cultural bias from intelligence tests, some psychologists have attempted to design culture fair tests. Such tests attempt to include only items to which groups regardless of social, racial and ethnic background have been exposed. These tests tend to be nonverbal in nature. However, these efforts have been partly successful and cultural bias has not been completely eliminated.

Current Debate on Intelligence Testing

What do intelligence tests really assess? Debate on the structure of intelligence is still hot without any satisfactory end. In recent years, Ceci (1991), Deary (1995) and many other psychologists have worked to identify the basic cognitive mechanisms and processes that underline intelligence and enable people to obtain high scores on intelligence tests. This has led to major developments: first, several intelligence tests based on a growing understanding of many aspects of cognition, have been developed (Naglieri, 1997). Among these the most noteworthy are the Kaufman Assessment Battery for Children and the Kaufman Adult intelligence Test (Kaufman & Kaufman, 1983), Woodcock-Johnson Test of Cognitive Abilities (1989). These tests attempt to measure important aspects of fluid and crystallized intelligence.

Second, a growing body of research has focused on the findings that the speed with which individuals perform simple perceptual and cognitive tasks is often correlated with scores on intelligence tests (Neisser et al., 1996; Vernon, 1987). Fry and Hale (1996) have reported that fluid intelligence is directly proportional to the processing speed with which one can process information. Research evidence indicates that as children grow older, their processing speed increases. These gains in processing speed underline improvements in working memory and this in turn, leads to gain in fluid intelligence.

One of the possible measures to estimate processing speed is reaction time. However, Deary and Stough (1996) argue that such measures are indeed correlated with scores on intelligence tests, but these correlations are not as strong as might be expected. Finally, because reaction time tests tap speed of responding (e.g., pushing one of the two buttons) as well as speed of mental processes, scores on these measures may be affected by factors unrelated to intelligence. These problems have led the researchers to a different kind of measures of the speed of mental operations i.e., *Inspection time*, which appears to be a very promising measure for probing the nature of human intelligence. The shorter the duration of time necessary for the individuals to attain a given level of accuracy, presumably, the faster the speed of important aspects of their cognitive (mental) operations.

Cross-Cultural Testing of Intelligence

Since the middle of 20th century, the testing of persons from different cultures has received increasing attention. The developing nations need tests for maximum utilization of their human resources such as for admission, educational and counselling purposes, job selection and placement of personnel according to their abilities. Similarly, the problems of social and political development of some developing countries have greatly stimulated cross-cultural testing. Some of the earliest cross-cultural tests were developed for the testing of the large wave of immigrants coming to United States at the turn of the nineteenth century. Other early tests originated in the basic research on the comparative abilities were for relatively isolated cultural groups.

For many years intelligence tests have been severely criticized. The main objection has been that such instruments do not really measure innate potential; rather they are loaded with cultural biases and impact of cultural parameters along which cultures vary. A well-known example of such a parameter is language. If the cultural groups to be tested spoke different languages, tests were developed that required no language. When educational background differed widely, reading was ruled out. Another parameter in which cultures and subcultures differ is of speed. Not only the tempo of daily life, but also the motivation to hurry and the value attached to rapid performance vary widely among national, cultural, and ethnic minority groups within a single nation and between urban and rural subcultures (Klineberg, 1928; Knapp, 1960; Womer, 1972). Still other parameters along which cultures differ pertain to test contents. Several non-language tests, for example, call for items of information that are specific to certain cultures. To minimize the influences of various cultural parameters, several noteworthy but largely unsuccessful attempts were made to develop *culture free* intelligence tests. Later on the objective was modified to construct *culture fair* tests, that may only include items related to experiences common to a wide range of cultures and eliminating certain parameters, such as reading, speed of responding etc. (Anastasi, 1997; & Urbina, 1997).

Approaches to Cross-cultural Testing

Three different approaches have been followed in construction of cross-cultural tests (Anastasi, 1997). The first approach involves the selection of items common to many cultures and validation of the resulting test against local criteria in many different cultures. This is the basic approach of the culture-fair tests, although their repeated validation in different cultures has often been either neglected altogether or inadequately executed without such a step. However, it is difficult to establish that a test is relatively free from the culturally restricted elements. Moreover, it is unlikely that any single test could be designed that would fully meet these requirements across a wide range of cultures.

The second approach is to develop a test within one culture and administer it to persons with different cultural backgrounds. However, any test developed within a particular cultural context can not be used as a universal yardstick for measuring intelligence. High and low scores on such a test may need different causal explanations when obtained by members of different cultural groups. Such an approach provides cultural distance between groups, as well as the individual's degree of acculturation and his/her readiness for educational and vocational activities that are culture-specific. From time to time, investigators have followed this approach in order to demonstrate the fact that cultural milieu in which an individual is reared affects the cognitive skills and knowledge s/he acquires.

As a third approach, different tests (or substantial adaptation of existing tests) may be developed within each culture, validated against the local criteria, and used only within appropriate culture. This approach is illustrated by the development of tests for industrial and military personnel within particular cultures. In this approach, the tests are validated against the specific educational and vocational criteria they are designed to predict, and performance is evaluated in terms of local norms. Each test is applied only within the culture in which it was developed and no cross-cultural comparison are attempted.

The two widely used tests that probably come as close as any to be culturally fair are the Raven's Progressive Matrices and Cattell's Culture-fair Intelligence Test. The Raven's Progressive Matrices (J. C. Raven, 1983; J. Raven, J. C. Raven, & Court, 1995) was designed primarily as a measure of Spearman's "g" factor. In keeping with Spearman's theoretical analysis of "g" factor, this test requires chiefly the eduction of relations among abstract items. The test consists of a set of matrices arranged into rows and columns, from each of which a part has been removed. The task is to choose the missing inserts from given alternatives. The easier items require accuracy of discrimination; the more difficult items involve analogies, permutation and alternation of pattern, and other logical relations. The test is usually administered with no time limit and can be given individually or in group. The Raven's Progressive Matrices (RPM) is available in three forms, differing in difficulty level. The *Standard Progressive Matrices*

(SPM-1996 edition), the *Coloured Progressive Matrices* (CPM-1990 edition) for young children and for special groups who can not be adequately tested with the SPM for various reasons, and the *Advanced Progressive Matrices* (APM-1994 edition), which has been developed for above average adolescents and adults. APM is used extensively with higher-ability adults and its popularity is partially attributed to its apparent low level of culture loading. Enhanced performance requires no special cultural knowledge, such as general information or vocabulary, because all its items pertain to geometric matrices. For this reason experts have argued that it is the purest available measure of g or analytical (fluid) intelligence (e.g., Carpenter et al., 1990; Jensen, 1980; Raven,1989). Several factorial analyses suggest that the Progressive Matrices are heavily loaded with a factor common to most of intelligence tests (identified with Spearman's "g" by British Psychologists), but that spatial aptitudes, inductive reasoning, perceptual accuracy and other group factors also influence performance.

Although the matrices are excellent items, the Raven's matrices on its own is limited simply because there is insufficient variety of items. Inevitably it will favour those high on the factor specific to the items and be biased against those who are low on this factor, which is of course, irrelevant for measuring intelligence (Kline, 1991).

The Culture-Fair Intelligence Test (Cattell & Cattell, 1957) is a paper and pencil test based on Cattell's theory of fluid intelligence. It is available in three levels: scale 1, for ages 4 and 8 and mentally retarded adults; scale 2, for ages 8 to 13 and average adults and scale 3, for grades 10 to 16 and superior adults. Each scale has been prepared in two parallel forms, A and B. Scale 1 requires individual administration for at least some of the tests; the other scales may be given either as individual or as group tests. Each scale is composed of 4 subtests for measuring individual's ability to perceive relationship among things: *Series, Classifications, Matrices* and *Conditions*. This test is specifically to be as culture fair as possible and to come as close as possible to measure "g_f" (fluid intelligence). The test includes four types of nonverbal items, which are equally familiar or unfamiliar to all cultures. In addition to 4 subtests, scale 1 contains four other subtests to measure cultural information and verbal comprehension. The entire scale takes 40-60 minutes working time. The Cattell's tests have been administered in several European countries, in North America and in certain African and Asian cultures. Norms tended to remain unchanged in cultures moderately similar to that in which the tests were developed. In other culture, however, performance fell considerably below the original norms. Moreover, black children of low socio-economic status tested in United States did no better on this test than on the Standford-Binet (Willard, 1968).

Problems in Cross-Cultural Testing

When psychologists began to develop instruments for cross-cultural testing in the first quarter of the twentieth century, they hoped it would be at least theoretically possible to measure *hereditary intellectual potentials* independently of the impact of cultural experiences. The individual behaviour was thought to be overlaid with the sort of cultural veneer whose penetration became the objective of what were then called *culture free* tests. Subsequent developments in genetics and psychology have demonstrated the fallacy of this concept. It is now recognized that heredity and environmental factors operate jointly at all stages in the organism's development and their effects are inextricably intertwined in the resulting behaviour. For human, culture permeates nearly all environmental contents. Since all behaviour is affected by the cultural surroundings in which the individual is reared and since the psychological tests are samples of behaviour, cultural influences are reflected in test performance. It is therefore, futile to devise a test that is free from cultural influences. The present objective in cross-cultural testing is rather to construct tests that presuppose only those experiences that are common to different cultures, for the reason, such terms as *culture-fair*, *cultural-common* and *cross-cultural* have replaced the earlier culture free level (Anastasi, & Urbina, 1997).

No single test can be universally applicable or equally fair to all cultures. There are as many varieties of culture-fair tests as there are parameters along which cultures differ. A non- reading test may be culture-fair in one situation, a non language test in another, a performance test in a third and a translation and adaptation of a verbal test in a fourth. The varieties of available culture-fair tests are not interchanged but are useful in different types of cross-cultural comparisons. It is also unlikely that any test can be equally fair to more than one cultural group, specially, if the cultures are dissimilar. While reducing cultural differentials in test performance, a cross-cultural test can not completely eliminate them. Every test tends to favour persons from the culture in which it was developed. The mere use of paper and pencil or the presentation of abstract tasks having no immediate practical significance will favour some cultural groups and handicap others. Emotional and motivational factors likewise influence test performance such as the intrinsic interest of the test content, rapport with the examiner, drive to do well on a test, desire to excel others and past habits of solving problems individually and cooperatively (Anastasi & Urbina, 1997).

Vernon (1969) presents a long list of problems involved in testing intelligence cross-culturally. Factors that affect the test scores may not reflect the true intelligence of the subjects. These include unfamiliarity with the test situation, lack of motivation, anxiety, excitement, and suspicion of tester. These problems are common when the psychologist is of a different race. In some cultures, there may be difficulties with particular types of items or materials. Lack of test sophistication adversely affects test performance and the negative effects can be even more severe in cross-cultural settings.

Warburton (1951) found that the form-board test, which requires the subjects to pass objects through a matching hole, was useless for the Gurkhas who simply used force to achieve the desired results. Clearly this was an invalid test for this sample. Pictures and diagrams too can create difficulties. This fact has been fully documented by Deregowski (1980).

A problem raised by Vernon (1979) concerns the handicaps caused by poor medical care and nutrition, which can lower intelligence test scores. These handicaps are restricted to cultures other than the West, where they are more prevalent. There may be other handicaps in some cultures. The most important environmental factors that Vernon (1979) lists as likely to affect intelligence test scores are: lack of varied perceptual and aesthetic experiences; restricted linguistic stimulation; lack of interest in formal education within family; restricted environmental stimulation; little schooling; emphasis on rote learning in school. Not all cultures will show all these deficits but, if present, these are likely to limit performance on intelligence tests in such a way that scores do not accurately reflect the potential ability of the subjects as they do in the West, where many of these handicaps are eliminated except among a small minority. It is also possible that besides variations in special aptitudes, there are genuine genetic differences between some cultural groups. For all these reasons to compare different cultural groups on an intelligence test and report the differences as reflecting differences in intelligence is highly simplistic. Any differences could simply reflect difference in some of the factors discussed above (Kline, 1991).

If a test is adapted in a particular culture so that items are meaningful, the tasks familiar and the instructions comprehensible and if the children or adults understand the nature of the test and its purpose, it is probably possible to test intelligence with some degree of accuracy. However, comparison with the results of other cultures will be still dubious (Kline, 1991). It should be pointed out, that if the above mentioned conditions are not satisfied, the test will not be valid. Validity must be demonstrated by factoring the test and ensuring that its structure is similar to its western original; and by correlating the scores with external criteria such as occupational success, school examination etc.

Does this mean that cross-cultural comparisons on tests are impossible? Many cross-cultural psychologists (e.g., Berry & Dasen, 1974) consider that this is indeed the case and it is not meaningful to compare the intelligence of members of different cultures. For the simple reasons that different qualities are valued in different cultures, crystallized intelligence may not be comparable at all.

Vernon (1979) suggests that comparison between two cultures might be possible if both the groups are equal regarding access to education, freedom from physical disabilities such as those arising from the malnutrition, familiarity with the test, freedom from the test anxiety, and valuation of the skills involved in the test. This argument rules out most cross-cultural comparisons, especially between developed and third world countries who could never meet these conditions. It is also interesting to note that these criteria rule out comparison between black and white children in south Africa and many might argue that this is also true of comparison between black and white in U.S.A. It appears that comparison on tests of intelligence between different cultures are dubious even though accurate measurement within cultures is possible if tests are properly developed.

Cattell (1971) has argued that his culture-fair tests, which use items that are nonverbal and equally unfamiliar to all subjects regardless of culture, enable cross-cultural comparison to be made of fluid ability. However, it does appear that familiarity with these material affects results, a familiarity which still differs between cultures even if it is not as obvious as with verbal or information-based items. Culture-fair tests have been claimed to be effective within cultures with subjects of differing educational background, but their results have to be treated with great caution (Kline, 1991).

Lynn (1987) claims that mongoloids are superior to the caucasian on Spearman's "g" and spatial ability but are lower in verbal abilities and in rate of maturation. These differences are attributed to evolutionary pressure on the mongoloids in the extreme cold of the Ice Age, which has resulted in neurological differences, with more of the mongoloids cortex being devoted to spatial than to verbal ability. Lynn, Pagtiari, and Chann (1988) report substantially similar results for studies of children in Hong Kong, where similar differences in mongoloids abilities were found i.e., they are best in spatial ability and g but low on verbal ability (Kline, 1991).

Test Development in Pakistan

The research studies in the development of tests described above were mostly conducted in Europe and America, however, some work on testing has also been done in other parts of the world as well. In the subcontinent, researches on development and adaptation of tests started before partition. As early as 1920s, work on adaptation of tests was initiated in India. Kamat in 1935, adapted Stanford-Binet Scale in Marathi and Kannada languages. In Northern India Dr. Rice had produced Urdu and Punjabi versions of this scale. In Pakistan many research studies have been conducted on development, adaptation and validation of intelligence, ability, aptitude, achievement, personality tests and other related issues. However, it should be mentioned here that in Pakistan, there is scarcity of research on the development of intelligence tests and particularly nonverbal intelligence tests. Although some efforts have been made to validate tests like Raven

Matrices, Cattell's Culture-Fair Intelligence Test and Wechsler Intelligence Scales for adult and children on different populations and against different criteria, but little work has been done in the field of development of an indigenous nonverbal test of intelligence, especially, for adolescents and adults. The only work, worth mentioning, in the field of nonverbal intelligence testing in Pakistan was carried out by Psychological Section at Army General Headquarters. In 1950, various types of nonverbal paper pencil and performance tests were designed on the pattern of different famous nonverbal and performance tests. They were standardized on a large population of students from a large number of schools and colleges from different parts of the country.

In Pakistan, intelligence tests are being used for personnel selection since independence. The major users of intelligence tests in the country are Armed Forces, Federal Public Service Commission and four Provincial Public Service Commissions. Recently intelligence tests have also been included in the newly introduced Entry Tests being used for selection of candidates for admission in Medical Colleges and Engineering Universities throughout Pakistan. Most of these tests were developed during World War 1 and introduced by British in the subcontinent. Since then little or no improvement has been made in culturally adapting them, testing their construct validity and determining the norms anchored in the cultural reality of our society. Perhaps very little efforts have been undertaken by the concerned psychologists to develop indigenous intelligence tests for personnel selection (Shah, 1994).

An indirect indication of the need to develop psychological tests in Pakistan was first made by the commission on National Education in 1959. The commission expressed dissatisfaction with the existing examination system and criticized its reliability and validity. It recommended the establishment of a Bureau for the development and construction of tests to develop and to use more objective methods of guidance and evaluation in education. Since then almost all the commissions and committees appointed by the government to recommend reforms in education have recommended the need for intelligence and aptitude tests and other objective examination methods.

In 1962, Board of Secondary Education Karachi established a project to develop and standardize general ability and aptitude tests for educational guidance at the secondary school level (Hassan, 1986). The senior teachers wrote items for the tests, which were standardized on a sample of 2000 children. However, the use of these tests was rather limited and was soon abandoned. Similarly, little use of aptitude tests developed in the former East Pakistan at Dacca university and at the Punjab university was made (Ansari, 1979). Further, there has been much duplication in research on adaptation and development of tests because of the lack of any centrally coordinating body at the national level e.g., National Testing Bureau. For instance, at least, four different institutions have worked on the development of aptitude tests, the adaptation of WISC and the translation and adaptation of Study Habit and Attitude Questionnaires (Ansari, 1979). There has been no effort at national level to establish norms to ascertain use of these scales across different subcultural and regional groups throughout Pakistan. Different institutions including universities and other organizations are using psychometric methods to develop/adapt tests according to their own requirements and available expertise. However, no serious attempt has been made so far to develop and standardize indigenous tests. A brief summary of some of the research studies conducted on the development of intelligence tests and related issues in Pakistan is presented below:

Jamal (1964) used Raven's Colour Progressive Matrices (CPM) with Kindergarten children to explore the relationship between socio-economic level and intelligence. The differences found were not significant.

Kausar (1998), carried out a research to study the validity of Raven's Coloured Progressive Matrices (CPM) in Pakistan. To study the latent intellectual ability of children under the age of 11years, a random sample of 324 students was drawn in terms of age, grade, gender, urban-rural residence, socioeconomic status and level of academic achievement. The results of the study revealed significant differences in performance on CPM in favour of urban school children, girls and high academic achievers. However, the effects of SES were not found statistically significant. The findings of the study suggest that CPM is a valid measure of latent intellectual ability for children belonging to urban areas of Pakistan. However, the test seems to be biased against the rural areas.

Ansari and Iftikhar (1984) conducted a study to determine the validity of Raven's Standard Progressive Matrices for urban and rural school children in Pakistan. It was found that RSPM is useful as a test of intellectual performance for the urban school children. For the rural school children the utility of this test was limited.

Ismail and Mehmood (1986) administered Raven's Standard Progressive Matrices to 300 students to study the effect of sex and social class. A significant difference was found among the performances of three different classes. However no significant difference was found between boys and girls.

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

Riaz (1979) conducted a study to find out distinction between the constructs, intelligence and creativity and their relationship with academic achievement. Multiple correlation of intelligence and creativity with academic achievement showed that addition of creativity test to intelligence adds significantly to the prediction of achievement over and above obtained by intelligence test alone.

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.

Sheikh, (1982) developed "Zahanat Paima" in Department of Applied Psychology, University of the Punjab. It is an adaptation of Otis Quick Scoring Mental Ability Test in Urdu. However, its use has been limited to the Department concerned and has never been available to other researchers and psychologists.

Ain (1985) validated Cattell's Culture-Fair Intelligence Test on Pakistani children. The researcher administered Scale II, form A of Culture Fair Intelligence Test to 1129 students of 5th, 6th and 7th grades of English and Urdu medium schools of Peshawar. The validity study demonstrated that although test scores were not significantly correlated with age, they were strongly related to school grades, academic achievement and teacher's ratings. The author explained low correlation of test scores with age in terms of non availability of accurate record of the age of these students.

Israr (1985) tried out thirteen Piagetian tasks on a sample of 360 primary school children from all over Pakistan. The result show that grade 1 children were at early concrete operational stage, grade 3 at mid concrete operational stage and grade 5 were at late concrete operational stage.

In a study at National Institute of Psychology (NIP), Khan and Ahmad (1984) developed a cultural adaptation of Columbia Mental Maturity Scale. It is meant for a lower age group (3 years 6 months to 9 years 11 months). It is a test of general reasoning containing geometrical and figural material, some of which are coloured. The major changes in adaptation were replacing colours with line design for ease of reproductions and changes in figural material to make them closer to the experiences of the Pakistani children. Geometrical drawings were retained as such.

Aziz (1997) developed a Pakistani version of Columbia Mental Maturity Scale (CMMS) for children aged 3 years 6 months to 10 years. This test consists of 92 pictorial and figural classification items. These items are arranged in a series of seven overlapping levels: two for pre-school children, and five for each grade 1 through 5. The reliability and validity indices characterized the test as a useful tool for children in Pakistan, specifically in the school setting.

. To investigate differential schooling effects on cognitive development and intellectual ability in the primary school students, Khalid (1986) studied differences in cognitive development of primary school students of government and private schools. The results revealed no significant correlation between school type and cognitive development indicating that the differences between cognitive levels of the children of different schools are largely due to the socioeconomic status of the parents and probably other home variables with negligible impact of the school.

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

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

Israr and Abbas (1990) developed a Test of Intellectual Development for pre-school children. The contents of the test are both verbal and nonverbal. It consists of eight subtests; *Colour naming, Reasoning, Seriation, Verbal-memory, Pictorial-memory, Perceptual-motor task, One to one correspondence* and *Conversation.* The reliability of the test was determined by test-retest and KR-20 methods. The test was validated against the criterion of age differentiation. The reliability and the validity studies confirmed the utility of the test as a sound psychometric tool.

Mahmud (1990) developed and validated Educational Ability Test for Pakistani pre-school children. The test consisted of 56 items covering six areas: *Visual Matching, Reasoning, School Language, Quantitative Concepts, Auditory Memory* and *Rhyming*. Test-retest reliability (.82) demonstrates temporal stability of the test over a period of two months. Similarly, reliability of the test computed by KR-20 formula (.90) shows the homogeneity of the test. Significant differences were found in the rural-urban samples, while gender differences were indicated only in rural sample.

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 a student comes across in his/her environment in the school and outside the school. The internal consistency and test-retest reliability were found to be satisfactory ranging from.87 to .90 for various groups.

Hussain (1992) developed a Group Verbal Intelligence Test in Urdu for high school students. The test comprised two subtests; *Vocabulary Test* and *Numerical Ability Test*. The reliability of the test was determined by KR-20 method. It was validated against the criterion of school marks. The reliability (.88) and validity (.55) were found to be significantly high. Significant differences were found in male and female sample.

Naheed (1993) has developed a Verbal Test of Intelligence for Pakistani Urban Primary School Children. The test is comprised of two sub-tests: *Vocabulary* and *Arithmetic*. The school marks were used as a validity criterion for the test. The reliability was determined by KR-20 formula. Both the reliability (.80 to .85) and Validity (.89 to .90) indices were found to be highly significant.

Syed (1993) developed a Non Verbal Test of Intelligence for Pakistani Urban Primary School Children. It is an individual test comprising two subtests: *Block Design* and *Picture Completion*. The reliability of the test was determined by KR-20 method while validity index was obtained by correlating the test scores with school marks of the subjects. The reliability (.82 to .86) and validity (.85) indices were found significantly highly.

Gardezi (1994) developed a Non Verbal Intelligence Test for students of grade 10 in the age range of 15 to 17 years. The test comprised four subtests viz., *Series, Analogies, Classification* and *Matrices*. Both the reliability (.77 to .82) and validity (.76 to .82) indices were found to be highly significant. Percentile norms were developed separately for boys and girls.

RATIONALE OF THE STUDY

RATIONALE OF THE STUDY

The study is aimed at developing and standardizing an Indigenous Nonverbal Test of Intelligence (INTI) for the assessment of general intelligence of Pakistani youth. The main emphasis is to develop a nonverbal test of general intelligence "g" in a manner to minimize the influence of verbal fluency, specific social and school learned skills, knowledge, socio-economic status, educational, cultural and experiential background on the test scores by using nonverbal test content comprising figurative material from our own cultural and folk heritage.

Like most of the developing countries, there is scarcity of research on test development in Pakistan. Most of our organizations and testing services are still using foreign tests without even proper adaptation and validation against local criteria. Although some efforts have been made in adapting tests imported from the west like WAIS, WISC, Raven's Progressive Matrices, Cattell's Culture Fair Intelligence Test etc. but no serious effort has been made to develop and standardize indigenous tests of intelligence. Majority of the newly developed/adapted tests are half baked, where loose psychometric criteria, and even faulty methodology have been applied which raises serious doubts about the reliability, validity and especially the construct validity of these tests. In most of these studies, the sampling is also a major limitation. The researchers have neither used national sample nor have included people from rural areas in the sample who constitute bulk of our population. Similarly due care has also not been taken to include subjects from all ethnic groups and social classes etc. Item analysis, and reliability and validity studies have been conducted improperly employing small samples. In most of the cases, the statistical analyses performed are inadequate. The use of foreign intelligence tests or their adapted versions with psychometric shortcomings in measuring intelligence to predict success in school/college or on the job makes the results questionable (Shah, 1994).

As regards the question, why to develop a new non verbal test of intelligence and why not to use the existing well established culture fair (nonverbal) intelligence tests like Raven Matrices and Cattell's Culture Fair Intelligence Test, many educators, psychologists and personnel officers of the developing countries hold widely divergent views. At the one extreme, there are those who look mainly at the vast environmental differences between the developing countries and the highly advanced nations and conclude that any test designed for one culture can not serve the other. At the other extreme, there are many others who attach greater importance to the fact that the skills needed in both developed and developing countries are exactly the same, so these tests can be used everywhere without adaptation (Schwarz & Krug, 1972). Still there is another point of view that these tests can be used cross-culturally after proper adaptation and validation against local criteria.

It has been observed from different cross-cultural studies that culture fair or cross-cultural intelligence tests are not universally applicable or equally fair to all cultures in view of their numerous dissimilarities. The Culture Fair or Cross-Cultural tests can not completely eliminate cultural differentials. Every test tends to favour persons from the culture in which it was developed. The mere use of paper and pencil or the presentation of abstract tasks having no immediate practical significance will favour some cultural groups and handicap other (Anastasi, & Urbina, 1997).

Intelligence is very closely related to cultural technology. In the West, the cultural technology is literacy. It is therefore, impossible to test intelligence in different cultures with the tests developed in the West unless we know more about their technologies. If the test is designed to provide a good description of competence in a literate society, it is unlikely to have anything in common with tests suitable in an illiterate or less literate society. Unless we take into account all fundamental operations that cut across all forms of intelligence and intelligent behavior, we can not construct tests for different cultures. Still, we have to have operations that test through culturally relevant items. For the Raven's Matrices, it should be more appropriate to construct a test in which the figures should relate to the patterns or designs commonly used in the culture in which the test is going to be used. Such patterns will have at least much higher face validity than the items used in the original test (Mahmood, 1991).

Sternberg (1985), in his triarchic theory of intelligence 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 and it cannot be understood and assessed outside socio-cultural context. What may be relevant in one culture may not be in another culture.

Cross-cultural nonverbal intelligence tests cannot be used effectively without proper adaptation and validation against local norms. Parmer (1989) in his study of cross-cultural transfer of nonverbal intelligence tests provides evidences that without adequate examination of validity of culture fair test in a particular culture, it does not demonstrate the expected equal performance across cultures. He suggests that exclusive use of abstract- figural test and test not standardized on relevant sample should be avoided.

After adapting a culture-fair/nonverbal test to a particular culture, or proper validation against local criteria, it is probably possible to test intelligence with some degree of accuracy. But still it is unlikely that any single test could be designed that would fully meet the requirements across the wide range of cultures. The problem of construct validity of the test is always there. The test material, for example, complex structure and geometrical figures of intelligence test that pertain to one cultural construct may be irrelevant in other cultures in their original forms. The results of such tests in other cultures always remain dubious (Kline, 1991; Shah, 1994).

There is a greater cultural diversity in the Third World countries than in the industrialized and developed countries of the West. Another problem in the adaptation of foreign tests is the strong urban and rural divide in the Third World countries. In a primarily rural society, parts of these countries become urbanized and industrialized and the development is often lopsided. Consequently, the social, economic and cultural differences between rural and urban populations are ever increasing with the passage of time. As a result, even the so-called culture-fair tests have been shown to differ significantly from the normative data of rural and urban populations. In an interesting and pioneering study (Ansari & Iftikhar, 1984), Raven's Standard Progressive Matrices were given to urban and rural school children of 6th, 7th and 8th grades. It was found that the gap between the two increased with age, urban pupils doing better than the rural ones, at higher age levels (Mahmood, 1991).

According to Brislin (1990), intelligence can best be understood and measured in cultural relative terms. One cannot perform well on an intelligence test unless s/he is familiar with the culture from which it has originated.

Cattell (1971) has claimed that Culture Fair Intelligence Tests (nonverbal in content) are more effective within culture with subjects of different educational and socio-economical background. He has also warned that cross-culturally use of these tests have to be treated with great cautions.

In a recent study in United States, Naglieri & Prewett (1999) found that non-fluent English speakers and children with receptive or expressive language impairments are at disadvantage when assessed with traditional intelligence tests. They suggested that for intelligence assessment such methods be used that place little or no emphasis on English language skills.

There have been commonly two approaches to assessing intelligence in situations where the individuals in question would be handicapped. First, traditional tests are used, using standard or adapted testing instructions. For example, the Wechsler Intelligence Scale for Children, WISC-III (Wechsler, 1991) and Stanford Binet Intelligence Scale, SB-IV (Thorndike, Hagen, & Sattler, 1986) are often employed, partly because they are purported to use some nonverbal assessment techniques. However, the nonverbal abilities assessed through the performance scale of the WISC and nonverbal reasoning scale of Stanford Binet include tasks that require language based strategies for successful completion, and both tests include culture-bound items. In particular, the SB-IV relies heavily on language ability throughout and therefore is totally inappropriate for use with the individuals from different cultural and language backgrounds or level of hearing or speech abilities (Sullivan & Burley, 1999).

The second approach is to use specialized intelligence tests designed to measure intellectual functioning through a system that does require language-based responding. Raven's Progressive Matrices (Raven, Court & Raven, 1985), the Test of Nonverbal Intelligence (Brown, Sherbenou, & Johnsen, 1990) and the Matrix Analogies Test-Expanded Form (Naglieri,1985) are a few of the more popular tests that are specifically developed to measure intelligence nonverbally. Each of these tests has been criticized for a variety of shortcomings, most important being the narrow range of abilities they assess (Sattler, 1992). The Raven's Progressive Matrices and Matrix Analogies Test are used to measure nonverbal cognitive abilities solely through the use of figural matrices and the Test of Nonverbal Intelligence assesses intelligence solely through problem solving. The problem with using these types of tests is that they do not assess intelligence in a comprehensive manner. Naglieri and Prewett (1999) believe that a nonverbal measure of intelligence should provide a more complete evaluation of the cognitive processing of individuals with hearing impairments, physical limitations, limited knowledge of the English language, and language/ communication disorders, as well as those of normal persons.

According to modern factor analysis (Cattell, 1971, 1987), general intelligence "g" is composed of fluid (spatial-mechanical) intelligence "g_f" and crystallized (verbal-mathematical) intelligence "g_c". Cattell regards fluid intelligence as being closely related to a person's innate capacity for intellectual performance. It depends more on biological and physiological influence than on formal schooling. On the other hand, crystallized intelligence is what a person has accomplished through the use of his/her fluid intelligence and what s/he has learned is based on social, cultural and formal school learning. Fluid intelligence reflects the level of intellectual competence associated with casual learning processes and is assessed by nonverbal intelligence tests (measures of abstract reasoning), whereas, crystallized intelligence reflects intellectual competence associated with formal national/international learning processes and is assessed by traditional verbal tests (measures of vocabulary, world knowledge etc). Fluid and crystallized abilities are highly correlated at an early age of 2 or 3 years. As children grow older and undergo different experiences at school and in the family, fluid ability and crystallized ability become less highly correlated. The bright and well-adjusted children exposed to rich environments both at school as well as home can invest most of their fluid ability in the crystallized skills of their culture. On the other hand, the equally bright children from homes where education is not valued and who attend schools providing poor quality of education will not be able to fully invest their fluid ability. The school performance of such children may be poor as compared to average children who invest all their abilities at a school (Kline, 1991).

Although, many psychologists believe that 'intelligence' is synonymous with fluid intelligence, but in countries where educational system is reasonably egalitarian, crystallized intelligence may appear very similar to fluid intelligence (Eysenck and Kamin, 1981). Certainly, in USA, UK and continental Europe, the g_f and g_c correlate quite highly, whereas, in most of developing and underdeveloped countries, the correlation between fluid and crystallized intelligence is very low, which may be possibly explained in terms of obvious disparities in the educational systems

According to Cattell's theory, individuals' crystallized intelligence is dependent on their socioeconomic status, cultural, social, experiential and educational environments. In other words different individuals may have different levels of crystallized intelligence owing to differences in socioeconomic status, cultural and educational backgrounds. Therefore, it seems imperative to assume that crystallized intelligence of individuals measured by verbal tests may not represent an appropriate method to assess intelligence in a culture where numerous disparities exist in the educational, social and economical background of the individuals.

Pakistan is a country with many subcultures and classes. People differ in their educational, social and economical background. Most of the people of our country are from rural areas and middle or lower socioeconomic classes. Most of them fail to provide their children a congenial or conducive environment for better education and intellectual development. There are lots of disparities in our educational system too. The standard of education in our educational institutions, except few higher category national and international institutions, is very low. Under these circumstances it does not seem appropriate to judge or compare the intellectual abilities of our youth on the basis of their performance on foreign made verbal tests of intelligence that are inherently culturally biased. To measure the true intelligence or general ability of our youth with a minimal influence of their social, economical, cultural, experiential and educational environment, it is imperative to develop a nonverbal test of intelligence consisting of material developed indigenously from our own cultural and folk heritage.

In view of the insufficient, inadequate and psychometrically improper research studies conducted in our country on test development, difficulties in the administration of culture fair intelligence tests and significant differences in the educational, cultural and socioeconomical backgrounds of our people, it seems more appropriate to develop indigenous nonverbal test of intelligence for our youth based on scientific psychometric principles. The test should provide adequate evaluation of the intellectual/cognitive processing of individuals with different educational, cultural, experiential and socioeconomical backgrounds.

Objectives of the Study

Main objectives of the present study were as follows:

1. To develop a psychometrically sound indigenous nonverbal test of intelligence based upon comprehensive theoretical foundations with high reliability and validity for Pakistani youth.
2. To develop an instrument for assessment and prediction of intellectual potential of our youth.
3. To ascertain utility of the new test for all segments of population at national level.
4. To develop test norms in order to facilitate meaningful interpretation of scores both for individuals as well as groups.

METHOD

METHOD

TEST DEVELOPMENT

Planning the Layout

Planning is an essential activity in all stages of a test construction project. Test planning encompasses all of the many and varied operations that go into producing a new test. So, a blueprint of the proposed Indigenous Nonverbal Test of Intelligence (INTI) was prepared in line with the objectives of the present study, indicating the test content, generation of different types of items, formation of subtests of the proposed test, preparation of the draft test, instructions to be given, tryouts and analysis of test items.

Development of the Test Items

The items of the test were generated in a multi-stage process. First of all, the available literature/research studies regarding intelligence, ability and cognitive development testing, were reviewed. It was followed by a thorough study of some well known intelligence tests, specifically nonverbal tests of intelligence like Raven's Progressive Matrices, Cattell's Culture Fair Intelligence Test, Wechsler's Adult Intelligence Scale, and Otis-Lennon Mental Ability Test. Secondly, a panel of six psychologists from the National Institute of Psychology and Armed Forces who had expertise in the field of psychological test construction were consulted and their technical advice/guidance was taken for the development of the proposed indigenous nonverbal test and its contents. The review of the intelligence tests as well as research studies conducted in the field of intelligence test development and discussion with experts helped in preparing the blueprint of the test. The test was thus planned in a manner so as to measure the various aspects of general intelligence including abstract reasoning, spatial relations, conceptual ability, accuracy of discrimination, eduction of relations and correlates.

After the finalization of general content areas of the nonverbal test, the types of items were specified. Items were then developed according to the specifications laid down in the blueprint of the proposed test, using figurative material taken from our own culture and folk heritage including different patterns, shapes, drawings, designs, diagrams, symbols, pictures of the parts of human body etc. The main purpose of the development of such items was to minimize the influence of individual's educational, social, cultural and experiential background, specific social and school learned skills and knowledge on the test performance.

The items of the test were developed indigenously. However, ideas about the types of items were borrowed from different famous and standardized intelligence tests such as WAIS, Otis-Lennon Mental Ability Test, Raven's Progressive Matrices, Cattell's Culture-Fair Intelligence Test and Differential Aptitude Test (Abstract Reasoning). Items were constructed in the framework of English script writing (from left to right) instead of Urdu script writing (from right to left). This decision was based on the outcome of an empirical study (pilot study). The researcher devised two sets of 25 similar nonverbal items. One set had 25 items in English and other had 25 items in Urdu script writing. Both these sets of items were administered to a sample of 50 subjects one after the other. The subjects were asked to comment which version of the test items they found easy and convenient to solve. They were also asked to give reasons to support their opinion. After the analysis of data, no significant difference between the scores of the subjects on both sets of test items was found. Most of the subjects were of the opinion that they found English version of test items easier and convenient to solve. They were also of the view that as they learn and solve mathematical problems in English throughout their academic career, therefore, they always like to solve problems in that way. In the light of these findings, it was decided to develop test items in the framework of English script writing.

As the Indigenous Nonverbal Test of Intelligence (INTI) was planned to be a test of fluid intelligence which according to Cattell (1971) is best tested either by items which all members of culture have over learned or by items with which all subjects regardless of education and background are equally unfamiliar (Kline, 1986). While devising the test items, an effort was made to construct items, which could be equally, unfamiliar/familiar to all the subjects with different social, cultural and educational backgrounds or represent experiences common to every member in a particular culture. Item content consisted of material, which was neither specific to a particular culture, class, geographical area nor was advantageous to any class or group. However, the assumption of equality of experiences is questionable and it is difficult to develop such nonverbal items, which are equally unfamiliar or familiar to all the subjects. Some items may be closer to the experiences of subjects from one educational and social background whereas, other may prove favourable to other segment of the population.

Format of the Test

The format selected for the test was polychotomous. This format contains multiple choice items. There are two parts in a multiple choice item; (1) the stem, which contains the question or problem and (2) the options, which make up a set of possible answers from which subjects have to choose the correct one. All the items developed had 4 to 6 options depending upon the specific characteristics of different types of items. All the items had only one correct answer. The multiple choice items not only provide objectivity of measure but also cover a wider range of content areas. They are usually proved to be more appropriate for presenting concepts in novel situations in the stem. It was expected that this would assure the objective nature of the nonverbal measure both in administration as well as scoring. For the development of items of nonverbal intelligence test, rules for construction of items chalked out by Thorndike and Hagen (1991) and guidelines given in the " Standards for Educational and Psychological Testing" (1985) developed by a joint committee of the AERA, APA, and NCME were followed.

Types of Test Items

An initial pool of 250 nonverbal items comprising five main categories was developed. Each category contained 50 items.

Series

In these items, the subject is presented with a series of figures, patterns, designs etc., and is asked to choose the next one in the series, which is a logical continuation of the series amongst the choices provided as answers. In series completion items, one must encode the terms, then infer the relation between each successive pair of figures or patterns and then apply this relation to generate a correct completion.

Matrices

The kind of inductive reasoning item that combines elements of analogy, series completion and classification problems is the matrix item. In matrix item, there are typically nine small squares or cells embedded in one large square. In each of the small square is a figural design that is part of several patterns. The patterns go horizontally across the matrix and vertically down the matrix. Usually, one of the cells of the matrix is blank, most typically, the cell at the lower right. The examinee's task is to figure out what figural design from answer alternatives ought to go in the empty cell in order to finish the various patterns i.e., what figure from the answer figures will complete the horizontal and vertical patterns of which it is a part. Matrix item is a double figure analogy in which two relations must be identified.

Analogies

In this type of items, the subject's task is to deduce relationship between the terms and apply it in the new case. The subject has to detect the relationship in the first two figures or patterns and to find out which one of the figures given as answer options exemplifies the similar relationship in the third figure. In analogies, the mental processes of encoding, inference, mapping, application and justification are necessary. Typical geometrical analogies involve additions, deletions and transformation of geometrical figures or portion of such figures and the subject's task is to figure out what these additions, deletions and transformation are.

Odd-one-out (Classification)

Classification items require essentially the same set of performance components as series completion and analogy problems. Like series completion and analogies, classification items can come in a variety of forms. One form, the Odd-one-out consists of set of terms, one of which does not belong to others. The examinee's task is to figure out which term i.e., figure, design, pattern does not belong to other figures, designs and patterns etc.

Similarities (Classification)

Another format sometimes used in understanding and solving classification items is somewhat different. It consists of four sets of two terms each. Proceeding the four sets of terms is a single term appearing by itself. The examinee's task is to indicate with which of the four sets of two terms the single term should appear. The performance components used to solve such problems are much the same as those used to solve series completion and analogies, although that are applied in a slightly different way. Here first one must encode the terms of the problems, next infer what is common to each of the pairs of two terms and then map the higher-order difference or differences between the pairs of lower-order relations inferred. These differences will be further used as the basis for deciding to which of the four categories the single term belongs. Finally one needs to apply what s/he have learned in order to determine in which category the single term belongs.

Formation of Subtests

The items developed pertaining to five different categories were then assigned to relevant subtest. The five subtests were: *Series*, *Matrices*, *Analogies*, *Odd-one-out* and *Similarities* depending upon the format of items.

The five types of items not only make the test as varied as possible, they also provide an opportunity to measure different aspects of one's intellectual functioning without stress on any one particular ability or skill. The other reason for the construction of five subtests including different types of items is to minimize the error of measurement.

All the five types of items involve functions like abstract reasoning, spatial relations, conceptual ability, eduction of relations and correlates, permutation and alternation of alternatives, and accuracy of discrimination.

Initial Editing of the Items

The items prepared in the initial stage were presented to five psychologists involved in research and development of psychological tests for their expert opinion and technical advice. In the light of their views/suggestions, changes and adjustments were made in the problem and answer figures of the test items.

Instructions for the Administration of Test

The medium of instructions in our educational institutions varies from pure Urdu medium, a mixture of both Urdu and English and pure English medium. Therefore, for the convenience of subjects of all categories of educational institutions, general instructions for the administration of the test and specific instructions for each subtest were written in both Urdu and English languages. In this regard, efforts were also made to write Instructions in very simple, clear and easy language. A separate answer sheet for the test was also devised.

In order to select the most satisfactory items, the first draft test was tried out on a representative sample of target population.

First Try Out

A pool of 250 items of five different types was tried out on a sample of 50 students of grade 12 at Federal Government Sir Syed College, Rawalpindi. The test was administered on two groups of 25 students each. The students were introduced to the purpose of the study and were given instructions for taking the test. They were encouraged to suggest changes, they considered appropriate, for the improvement of the test, test items and answer sheet. The main objectives of the first try out were to:

1. test the adequacy of items and feasibility of test format.
2. discard items with ambiguous and unclear designs and drawings.
3. identify items, which were too easy or too difficult (low discriminatory power).
4. determine response valence.
5. arrange the items in ascending order of difficulty.
6. test the clarity and comprehension of general and specific task related instructions.
7. check appropriateness of the answer sheet.

On the basis of first try out, a total of 138 items out of 250 items, 30 each in first three subtests and 24 each in fourth and fifth subtest were retained. The remaining Items, which were reported by the subjects either too easy or too difficult, or having ambiguous and unclear drawings, designs etc., were discarded. Some modifications and improvements in the test format, test items and test instructions were also made. The response options of items, which were not checked by any respondent, were discarded or modified.

Second Try Out

After first try out, the format of the test and test items, instructions for the administration of test and test answer sheet were finalized. The second draft of the proposed test was administered to a representative sample of 200 subjects. It consists of 138 items grouped into five subtests. The *Series*, *Matrices*, and *Analogies* tests consist of 30 items each, while *Odd-one-out;* and *Similarities* contain 24 items each.

Objectives

The second try out was carried out for the following purposes:

1. Item analysis of the test with reference to estimating its:
    (a) Internal consistency
    (b) Item discrimination
    (c) Item difficulty
    (d) Response valence
2. Revision, improvement and arrangement of items and their response options in the respective subtest in the light of item analysis.

Sample

The second draft of the test was administered to a group of 200 candidates who had passed HSSC (grade12) and reported to Army Selection and Recruitment Centres for preliminary examination for commission in Pakistan Army. Their ages ranged from 17 to 22 years. The sample included subjects from different socio-economic, urban and rural and educational backgrounds. The list of the Army Selection and Recruitment Centers from where the sample was taken is given at annexure A.

Test Administration

The second draft of the test, comprising 5 subtests, was administered to subjects in groups. Each subtest was administered separately. Before the administration of test, efforts were made to fulfill the prerequisites of testing as far as possible. The subjects were seated comfortably in a relaxed situation to reduce their test anxiety. They were apprised of the objectives of the test and convinced that the test was only meant for research purposes and it had nothing to do with their preliminary examination for the Army. They were also assured of the confidentiality of results. At the start of the test, instructions for taking the test were given. Each of the subtest was explained with the help of examples. There was no time limit for taking the test at that stage, however, the subjects were asked to attempt all the items and complete the test as quickly as possible.

Scoring

The answer sheets of the test were scored with the help of hand scoring keys. The responses were coded in terms of one and zero, since each item had only one correct answer. The maximum scores on subtests I, II, III, IV and V could be 22, 26, 25, 22 and 19, whereas the minimum scores could be 4, 6, 6, 4 and 2 respectively.

Analysis of the Data

A computerized plan for the analysis of the items of the test was prepared and all the data were fed to computer. Since the main objective of the second try out was to determine the various psychometric properties of the test, therefore, the data were analyzed at different levels.

Item Analysis

As validity and reliability of any test depend ultimately on the characteristics of its items, each item of the draft test was analyzed quantitatively to compute three different indices: internal consistency, discrimination index and difficulty level.

Internal Consistency

To determine the homogeneity of each subtest, indices of internal consistency based on item-total correlation were computed. Consequently, only those items were retained that were internally consistent (p <.01).

Discrimination Index

The discrimination index (D) of each item was computed by contrasted group technique. For this purpose, 27% high scorers (U) and 27% low scorers (L) on each subtest were taken (Thorndike & Hagen, 1991). The discrimination index of each item was obtained by subtracting the number of persons answering it correctly in the L group from the number answering it correctly in the U group. The difference between U and L for each item was converted into proportion. The items having discrimination value less than .3 were eliminated assuming that these items failed to discriminate significantly between high and low scorers on the test. Two items were discarded because of negative discrimination indices. Such items were considered poor items as low scorers performed better than the high scorers on these items.

Difficulty Level

Difficulty level of an item refers to percentages of examinees attempting any item correctly. Percentages of responses to correct choices (p values) were calculated for this purpose. Difficulty level of each item was calculated in term of each subtest of the test. Items with difficulty level (p value) below .3 and above .7 were discarded. However, in each subtest one easy item was retained to establish rapport. Details of indices of internal consistency, discrimination and difficulty of each item of test are given in Table 1.

Table 1

Indices of Internal Consistency, Discrimination and Difficulty of the items of test

Items	Internal Consistency	Discrimination Index	Difficulty Level
Series			
1.	.40***	.35	.82
2.	.33***	.49	.57
3.	.42***	.66	.52
4.	.38***	.59	.50
5.	.44***	.54	.42
6.	.36***	.58	.64
7.	.43***	.51	.68
8.	.26***	.31	.72
9.	.22**	.30	.32
10.	.14	.09	.16
11.	.45***	.62	.53
12.	.31***	.41	.60
13.	.29***	.47	.54
14.	.50***	.67	.38
15.	.32***	.26	.80
16.	-.00	-.02	.08
17.	.13	.06	.08
18.	.17*	.17	.27
19.	.44***	.44	.69
20.	.35***	.58	.56
21.	.44***	.53	.30
22.	.33***	.64	.51
23.	.09	.02	.03
24.	.01	.02	.19
25.	.19*	.28	.31
26.	.28***	.51	.45
27.	.35***	.44	.31

Continue...

Items	Internal Consistency	Discrimination Index	Difficulty Level
28.	.44***	.52	.30
29.	.33***	.41	.31
30.	.33***	.41	.33
Matrices			
1.	.20*	.15	.89
2.	.36***	.51	.35
3.	.53***	.61	.38
4.	.14	.22	.61
5.	.50***	.62	.66
6.	.38***	.57	.41
7.	.25***	.22	.68
8.	.25***	.31	.62
9.	.16*	.20	.43
10	.27***	.27	.68
11	.39***	.52	.65
12	.41***	.51	.73
13	.42***	.58	.60
14	.49***	.63	.59
15	.49***	.68	.54
16	.49***	.67	.57
17	.01	- .02	.23
18	.37***	.59	.58
19	.47***	.67	.54
20	.38***	.52	.62
21	.36***	.50	.66
22	.37***	.54	.45
23	.58***	.76	.56
24.	.45***	.62	.62
25.	.19*	.20	.27
26.	.47***	.65	.53

Continue...

Items	Internal Consistency	Discrimination Index	Difficulty Level
27.	.27***	.27	.71
28.	.35***	.41	.31
29.	.35***	.41	.31
30.	.37***	.52	.34
Analogies			
1.	.15	.05	.92
2.	.13	.17	.38
3.	.41***	.51	.72
4.	.28***	.39	.69
5.	.13	.15	.50
6.	.22**	.30	.38
7.	.32***	.20	.77
8.	.42***	.67	.47
9.	.05	.00	.18
10.	.39***	.53	.70
11.	.33***	.26	.78
12.	.35***	.56	.44
13.	.35***	.46	.60
14.	.29***	.38	.63
15.	.38***	.58	.57
16.	.47***	.61	.33
17.	.24**	.37	.64
18.	.18*	.15	.26
19.	.02	.09	.23
20.	.26***	.41	.30
21.	.36***	.58	.59
22.	.48***	.54	.63
23.	.42***	.52	.60
24.	.52***	.72	.50
25.	.28***	.51	.41

Continue...

Items	Internal Consistency	Discrimination Index	Difficulty Level
26.	.45***	.46	.75
27.	.03	.00	.13
28.	.49***	.67	.44
29.	.49***	.66	.55
30.	.28***	.34	.58
Odd-one-out			
1.	.38***	.43	.66
2.	.10	.04	.31
3.	.10	.09	.16
4.	.42***	.65	.50
5.	.51***	.69	.55
6.	.33***	.38	.35
7.	.22**	.31	.30
8.	.48***	.65	.44
9.	.43***	.43	.70
10.	.45***	.67	.45
11.	.16*	.15	.17
12.	.44***	.44	.72
13.	.37***	.49	.60
14.	.47***	.57	.58
15.	.41***	.69	.46
16.	.49***	.51	.72
17.	.32***	.58	.42
18.	.39***	.48	.65
19.	.37***	.52	.39
20.	.17*	.20	.43
21.	.43***	.51	.62
22.	.43***	.52	.52
23.	.53***	.72	.52
24.	.36***	.54	.40

Continue...

Items	Internal Consistency	Discrimination Index	Difficulty Level
Similarities			
1.	.13	.22	.37
2.	.27***	.33	.64
3.	.42***	.60	.48
4.	.36***	.59	.43
5.	.09	.06	.16
6.	.10	.09	.29
7.	.15	.13	.12
8.	.20*	.22	.24
9.	.37***	.41	.70
10.	.43***	.42	.73
11.	.30***	.41	.62
12.	.38***	.58	.45
13.	.20*	.22	.18
14.	.47***	.51	.52
15.	.25***	.17	.14
16.	.62***	.80	.44
17.	.41***	.42	.35
18.	.00	.00	.11
19.	.28***	.31	.32
20.	.56***	.72	.51
21.	.39***	.33	.30
22.	.24***	.20	.16
23.	.58***	.73	.42
24.	.36***	.57	.38

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

After various analyses, a total of 48 items were discarded; 41could not fulfill the three-stage evaluation criteria i.e., appropriate item difficulty level, discrimination power and internal consistency. The remaining 7 items were found qualitatively unsatisfactory (unclear drawings/designs), though fulfilling selection criteria. The details of the items discarded from each subtest of the test are given in Table 2.

Table 2

The index number of items discarded from each subtest of INTI

INTI & Subtests	Items Discarded	Total
Series	1,10,15,16,17,18,23,24,25,30	10
Matrices	1,4,7,8,9,10,17,22,25,27	10
Analogies	1,2,5,7,9,11,18,19,26,27	10
Odd-one-out	2,3,11,12,14,18,19,20,22	9
Similarities	1,5,6,7,8,13,15,18,22	9
Full Test		48

Response Valence

The items qualifying the evaluation criteria were further put to response analysis i.e., to determine valence of each option for an item. Percentages of responses to each keyed response as well as each option were calculated. There was no item choice, which failed to attract subjects (zero valence). However, a total of 30 item choices failed to attract considerable number of subjects (Table 3). Upon qualitative analysis, these options were found confusing and unclear and were improved, modified or replaced accordingly.

Table 3

Item-wise list of response choices with lower valence

Subtests	Item No	Serial position of Choices with lower valence
Series		
	7	2
	9	2
	21	4
	26	1, 5
	28	5
Matrices		
	3	1
	19	3, 6
	20	4
	21	1, 5
	24	2, 3
Analogies		
	3	1
	4	3
	13	3
	20	4
Odd-one-out		
	1	4, 5

Continue...

Subtests	Item No	Serial position of Choices with lower valence
Similarities		
	3	2
	4	1
	12	1
	16	1
	19	2
	20	1
	21	3
	23	3,4
	24	2

Rearrangement of the Test Items

The remaining items in each subtest that qualified the evaluation criteria were rearranged in an ascending order of difficulty.

The Final Test

The final draft of the test consists of 90 items. Three subtests viz., *Series*, *Matrices* and *Analogies* are comprised of 20 items each whereas, two subtests, *Odd-one-out* and *Similarities* are comprised of 15 items each. Details of indices of internal consistency, discrimination and difficulty of each item of the final test are given in table 4. Distribution of items of the final test in different ranges of discrimination and difficulty levels is given in Table 5.

Table 4

Indices of Internal Consistency, Discrimination and Difficulty of the items of the final test

Items	Internal Consistency	Discrimination Index	Difficulty Level
Series			
1.	.26***	.31	.72
2.	.44***	.44	.69
3.	.43***	.51	.68
4.	.36***	.58	.64
5.	.31***	.41	.60
6.	.33***	.49	.57
7.	.35***	.58	.56
8.	.29***	.47	.54
9.	.45***	.62	.53
10.	.42***	.66	.52
11.	.33***	.64	.51
12.	.38***	.59	.50
13.	.28***	.51	.45
14.	.44***	.54	.42
15.	.50***	.67	.38
16.	.22**	.30	.32
17.	.33***	.41	.31
18.	.35***	.44	.31
19.	.44***	.53	.30
20.	.44***	.52	.30
Matrices			
1.	.41***	.51	.73
2.	.50***	.62	.66
3.	.36***	.50	.66
4.	.39***	.52	.65

Continue...

Items	Internal Consistency	Discrimination Index	Difficulty Level
5.	.38***	.52	.62
6.	.45***	.62	.62
7.	.42***	.58	.60
8.	.49***	.63	.59
9.	.37***	.59	.58
10.	.49***	.67	.57
11.	.58***	.76	.56
12.	.49***	.68	.54
13.	.47***	.67	.54
14.	.47***	.65	.53
15.	.38***	.57	.41
16.	.53***	.61	.38
17.	.36***	.51	.35
18.	.37***	.52	.34
19.	.35***	.41	.31
20.	.35***	.41	.31
Analogies			
1.	.41***	.51	.72
2.	.39***	.53	.70
3.	.28***	.39	.69
4.	.24**	.37	.64
5.	.48***	.54	.63
6.	.29***	.38	.63
7.	.42***	.52	.60
8.	.35***	.46	.60
9.	.36***	.58	.59
10.	.28***	.34	.58
11.	.38***	.58	.57
12.	.49***	.66	.55

Continue...

Items	Internal Consistency	Discrimination Index	Difficulty Level
13.	.52***	.72	.50
14.	.42***	.67	.47
15.	.49***	.67	.44
16.	.35***	.56	.44
17.	.28***	.51	.41
18.	.22**	.30	.38
19.	.47***	.61	.33
20.	.26***	.41	.30
Odd-one-out			
1.	.49***	.51	.72
2.	.43***	.43	.70
3.	.38***	.43	.66
4.	.43***	.51	.62
5.	.37***	.49	.60
6.	.51***	.69	.55
7.	.53***	.72	.52
8.	.42***	.65	.50
9.	.41***	.69	.46
10.	.45***	.67	.45
11.	.48***	.65	.44
12.	.32***	.58	.42
13.	.36***	.54	.40
14.	.33***	.38	.35
15.	.22**	.31	.30
Similarities			
1.	.43***	.42	.73
2.	.37***	.41	.70
3.	.27***	.33	.64
4.	.30***	.41	.62
5.	.47***	.51	.52

Continue...

Items	Internal Consistency	Discrimination Index	Difficulty Level
6.	.56***	.72	.51
7.	.42***	.60	.48
8.	.38***	.58	.45
9.	.62***	.80	.44
10.	.36***	.59	.43
11.	.58***	.73	.42
12.	.36***	.57	.38
13.	.41***	.62	.35
14.	.28***	.31	.32
15.	.39***	.33	.30

**p <.05, ** p<.01, *** p <.001*

Table 5

Distribution of items of INTI in different ranges of discrimination and difficulty levels

p Values	Number of Items	
	Discrimination	Difficulty
.71- .80	6	4
.61- .70	22	20
.51- .60	33	26
.41- .50	17	19
.31- .40	10	16
.21 .30	2	5
Total	90	90

Table 5 shows that INTI has a fairly normal distribution of items in terms of discrimination and difficulty levels. The highest number of items fall in the middle ranges of discrimination and difficulty (.41-.60).

Test Taking Time for the Test

After the final selection and arrangement of items in each subtest, a separate study was carried out to determine the average test taking time for each subtest and the full test. In this study, the test was administered to a sample of 100 subjects selected from the candidates who appeared for preliminary examination for commission in Army at Army Selection and Recruitment Centres. The subjects were told that it is a test of general ability. Each subtest was administered separately. The subjects were asked to complete the test as quickly as possible and hand over the answer sheets immediately after completion of the test. Time taken by each subject to complete each subtest was noted down on the answer sheet.

Test taking time for each subtest was determined by calculating the average time taken by first 85% of the subjects who completed the test. Average time for the full test was calculated by adding the average time taken for all the five subtests. The average time for each subtest and the full test is given in Table 6.

Table 6

Average time for each Subtest and the Full Test

INTI & Subtests	Number of Items	Average Time (Minutes)
Series	20	10
Matrices	20	10
Analogies	20	10
Odd-one-out	15	7
Similarities	15	8
Full Test	90	45

Final Study

Once the test and the time limit were finalized, the final draft of the test was subjected to another study to determine its different psychometric characteristics.

Sample

The sample consisted of 200 subjects, 100 urban and 100 rural. Their ages were between 17 to 22 years. The sample was selected from candidates who reported at different Army Selection and Recruitment Centres for preliminary tests for commission in Pakistan Army. As the basic qualification for commission in Army is Higher Secondary School Certificate (HSSC), majority of the subjects included in the sample had passed HSSC (grade 12) examination. However, 16 subjects who had taken their final examination of HSSC and were waiting for the results and 45 subjects who had passed HSSC, a year or two ago and were undergraduate students at the time of testing were included in the sample. The subjects comprising sample of the present study come from Urdu/English medium institutions, Science/Arts groups and different socioeconomic backgrounds. List of Army Selection and Recruitment Centres is given at annexure-B. Distribution of sample in terms of area (Urban/Rural) and medium of instruction (Urdu/English) is given in Table 7.

Table 7

Distribution of Sample in terms of Area (Urban/Rural) and Medium of Instruction (Urdu/English)

Medium of Instruction	Urban	Rural	Total
Urdu	48	64	112
English	52	36	88
Total	100	100	200

Test Administration

The procedure of test administration in the final study was the same as in the second try out except that in the final study, a time limit was set for each subtest. In this study the subjects were tested in groups and each subtest was administered separately. The subjects were told about the time limit at the start of each subtest. They were asked to answer the questions as quickly as possible and try to complete the test within the prescribed time.

RESULTS

RESULTS

Following statistical analyses were carried out to determine different psychometric characteristics of the test.

1. Item discrimination analysis (comparison between performances of high and low scoring groups on the test)
2. Mean, Standard Deviation and Mean percentages of correct responses of each subtest and the full test
3. Reliability of the test
4. Validity of the test
5. Inter correlations of the subtests

Item Discrimination Analysis

Item discrimination analysis is a comparison of performance of high scoring group and low scoring group on each item of the final test.

In order to recheck discrimination power of items of the final test, a chi square method was used. Two groups were formed, one consisting of 27% (n = 54) high scoring and the second 27% (n = 54) low scoring subjects on the test. Frequency comparisons were made on each item of the test between high and low scorers and their results (pass/fail). A 2 x 2 chi-square analysis was run for each item.

Table 8

Chi-square values showing discrimination power of the items comprising Indigenous Nonverbal Test of Intelligence (contrasted group technique)

Items	High Group (n = 54)		Low Group (n = 54)		
	Pass	Fail	Pass	Fail	Chi-Square Values
Series					
1	36	18	22	32	7.29**
2	47	7	34	20	8.34**
3	49	5	28	26	19.95***
4	37	17	20	34	10.73***
5	38	16	23	31	8.47**
6	37	17	19	35	12.01***
7	42	12	23	31	13.94***
8	36	18	23	31	6.31**
9	35	19	11	43	21.81***
10	45	9	23	31	19.21***
11	39	15	16	38	19.59***
12	32	22	14	40	12.26***
13	39	15	19	35	14.89***
14	33	21	14	40	13.59***
15	32	22	11	43	17.04***

Continue...

Items	Pass	Fail	Pass	Fail	Chi-Square Values
16	24	30	13	41	4.97*
17	37	17	19	35	12.01***
18	27	27	10	44	11.88***
19	31	23	10	44	17.33***
20	29	25	7	47	14.38***
Matrices					
1	43	11	23	31	15.58***
2	45	9	23	31	19.21***
3	46	8	27	27	15.25***
4	44	10	29	25	9.99***
5	42	12	24	30	12.62***
6	46	8	27	27	15.25***
7	39	15	15	39	21.33***
8	37	17	19	35	12.01***
9	38	16	12	42	25.17***
10	38	16	10	44	29.40***
11	39	15	12	42	27.08***
12	35	19	11	43	21.81***
13	28	26	8	46	16.66***
14	38	16	15	39	19.59***
15	30	24	12	42	12.26***

Continue...

Items	Pass	Fail	Pass	Fail	Chi-Square Values
16	35	19	2	52	17.08***
17	31	23	12	42	13.94***
18	27	27	11	43	10.39***
19	26	28	11	43	9.25**
20	24	30	13	41	4.97*
Analogies					
1	46	8	21	33	24.57***
2	52	2	35	19	17.08***
3	41	13	24	30	11.16***
4	43	11	29	25	8.16**
5	50	4	32	22	16.41***
6	41	13	26	28	8.84**
7	34	20	15	39	13.48***
8	33	21	13	41	15.14***
9	43	11	20	34	20.15***
10	39	15	23	31	9.69**
11	34	20	14	40	15.00***
12	46	8	22	32	22.87***
13	34	20	10	44	22.09***
14	33	21	9	45	22.44***

Continue...

Items	Pass	Fail	Pass	Fail	Chi-Square Values
15	41	13	12	42	31.15***
16	27	27	9	45	13.50***
17	35	19	13	41	18.15***
18	17	37	6	48	6.68**
19	33	21	10	44	20.44***
20	22	32	8	46	9 .04**
Odd-one-out					
1	33	21	13	41	15.14***
2	51	3	35	19	14.61***
3	42	12	24	30	12.62***
4	52	2	35	19	17.08***
5	42	12	24	30	12.62***
6	48	6	25	29	22.36***
7	51	3	23	31	33.65***
8	46	8	20	34	26.33***
9	46	8	22	32	22.87***
10	40	14	13	41	27.00***
11	41	13	16	38	23.22***
12	33	21	12	42	16.80***
13	31	23	10	44	17.33***

Continue...

Items	Pass	Fail	Pass	Fail	Chi-Square Values
14	28	26	11	43	11.59***
15	22	32	10	44	6.39**
Similarities					
1	49	5	32	22	14.27***
2	52	2	37	17	14.37***
3	47	7	35	19	7.29**
4	42	12	24	30	12.62**
5	38	16	19	35	13.41***
6	33	21	10	44	20.44***
7	27	27	5	49	21.49***
8	35	19	17	37	12.01***
9	51	3	23	31	33.65***
10	34	20	15	39	13.48***
11	34	20	5	49	33.75***
12	28	26	11	43	11.59***
13	46	8	22	32	22.87***
14	35	19	19	35	9.48**
15	20	34	7	47	8.34**

df =1, * $p<.05$, ** $p<.01$, *** $p<.001$

The results presented in Table 8 show significant differences in the performance of both high and low scoring groups on all the items of the test.

Table 9

Means, Standard Deviations and Mean percentages of correct responses of each Subtest and the Full Test (N = 200)

INTI & Subtests	Mean	Standard Deviation	Mean % of correct responses
Series	9.82	2.80	49.12
Matrices	9.55	3.08	51.50
Analogies	10.13	3.02	50.65
Odd-one-out	8.66	2.58	57.77
Similarities	7.72	2.18	47.78
Full Test	45.90	12.14	51.00

Table 9 presents the Means, Standard deviations and Mean Percentages of correct responses of five subtests and the full test. The Mean Percentages of correct responses reflect that subtest Odd-one-out is the easiest one whereas, Similarities is the most difficult one.

Reliability of the Test

Reliability is one of the major indices of the efficiency of any measure. The extent, to which one can depend upon a test, is very much determined by the reliability of the test. Following methods were used to establish test reliability:

1. Kuder Richardson Method (KR-20)
2. Split-half Method
3. Test-retest Method

Kuder Richardson Reliability

The inter-item consistency of the test was estimated by applying the Kuder Richardson Formula-20. Table 10 shows the KR-20 Reliability estimates for the subtests and the full test.

Table 10

Kuder Richardson Reliability of Indigenous Nonverbal Test of Intelligence (INTI)

INTI & Subtests	KR-20
Series	.59
Matrices	.64
Analogies	.63
Odd-one-out	.65
Similarities	.79
Full Test	.89

df =198

The results presented in Table 10 indicate high reliability of the test and homogeneity of the test items.

Split-half Reliability

To determine the split-half reliability, the test was divided into two halves based on odd-even items. Pearson product-moment coefficient of correlation was computed for the two half scores. The split-half reliability of the test is 0.74 ($p < .001$). In order to estimate the reliability of the full test, Spearman-Brown Prophecy formula was applied which yielded an increased correlation ($r = 0.85, p < .001$).

Test- Retest Reliability

Since stability over time is a very important consideration both for author as well as test user, it was decided to have an estimate of the temporal stability of the test. For this purpose, the test was administered twice with an interval of 30 days, on a sample of 100 students of grade 12 (other than the sample of the main study) selected from Army Public School and College for Boys, Rawalpindi. The two sets of scores thus obtained were used to calculate a coefficient of correlation indicating the reliability of the test. Table 11 gives the test-retest reliability coefficients of all the five subtests as well as the full test.

Table 11

Test-Retest Reliability Coefficients of Indigenous Nonverbal Test of Intelligence (INTI) with an interval of 30 days

	First Testing		Second Testing		
INTI & Subtests	*M*	*SD*	*M*	*SD*	Correlations
Series	11.42	3.11	11.99	2.80	.88***
Matrices	11.36	2.95	12.25	2.85	.82***
Analogies	11.89	3.20	12.64	2.89	.83***
Odd-one-out	9.10	2.52	9.75	2.29	.86***
Similarities	8.59	2.19	8.75	2.20	.81***
Full Test	52.36	12.43	55.30	11.63	.90***

*** $p < .001$

The test-retest reliability coefficients of INTI and its subtests are highly significant ($p <$. 001). These findings demonstrate that INTI is a reliable measure of intelligence.

Validity of the Test

Validity is an essential characteristic of any test. It refers to the degree or extent to which a test measures what it purports to measure. The validity of the test was determined by the following methods.

1. Grade/age Differentiation
2. Construct Validation
3. Criterion-related Validation

Grade/Age Differentiation

A major criterion employed in the validation of a number of intelligence tests is age differentiation. Such tests as the Stanford-Binet and most preschool tests are checked against chronological age to determine whether the scores show a progressive increase with advancing age (Anastasi & Urbina, 1997). Since cognitive abilities are expected to develop with age, it is expected that the mean scores on an intelligence test should likewise increase, if the test is valid.

Sometimes another similar criterion is used to establish the validity of an ability test, known as grade differentiation. During the developmental stage, it is expected that children in different school grades will show considerable variability in performance on ability tests. To determine the validity of the Indigenous Nonverbal Test of Intelligence, one of the validity criterion used in the present study was grade differentiation. For this purpose, mean scores of grades 10 to 12 students obtained on INTI were compared. These scores were derived from a sample (other than the sample of main study) of 210 students (70 students from each grade) selected at random from Army Public School and College for Boys, Rawalpindi. ANOVA was applied to determine the significance of differences in the mean scores of the three groups on all the subtests as well as the full test. All these findings are presented in Table 12. The data clearly indicates that mean ages of children in grades 10-12 show variations.

Table 12

Grade-wise differences in scores on Indigenous Nonverbal Test of Intelligence (INTI)

INTI & Subtests	Grade 10 (15 years & 8 months)		Grade 11 (16 years & 9 months)		Grade 12 (17 years & 10 months)			
	M	*SD*	*M*	*SD*	*M*	*SD*	*F*	*p*
Series	7.93	2.13	8.84	1.92	9.54	2.73	8.7	.000
Matrices	9.39	2.24	9.41	2.07	10.06	2.91	8.38	.000
Analogies	8.33	2.14	9.14	1.86	10.81	3.27	18.03	.000
Odd-one-out	7.4	2.11	7.82	1.39	8.51	2.29	8.01	.000
Similarities	6.24	1.85	6.43	1.38	6.89	2.48	2.45	.117
Full Test	37.29	8.87	40.85	7.77	45.87	13.06	9.47	.000

As evident from Table 12, all the F values are highly significant except for *Similarities*. Mean scores of INTI and all the subtests except *Similarities* show significant progressive increase with advancing grade which demonstrate validity of INTI. A further analysis of results based on mean age of the students of grade 10 to 12 also supports these findings, suggesting that test scores do provide empirical evidence that INTI can differentiate well between age groups.

Construct Validation

The construct validity of a test is the extent to which the test actually measures the theoretical construct or trait under investigation. Construct validation requires gradual accumulation of information from a variety of sources. Any data throwing light on the nature of the trait under consideration and the conditions affecting its development and manifestations represent appropriate evidence of construct validation. To determine the construct validity of INTI, the following techniques were used:

1. Factor analysis
2. Convergent and Discriminant validation.

Factor Analysis

Factor analysis is one of the important and relevant methods to construct validation procedures in identifying psychological traits. Essentially, it is a refined statistical technique for analyzing the interrelationships of behavior data. The factor analysis of INTI was carried out to explore if the newly devised test is a measure of fluid intelligence (the main objective of the present study). For this purpose the computer program SPSS, Inc. (1990) was applied. The data was subjected to principal component analysis. Results show high positive correlations among all the five subtests (Table 17). Principal Axis Factoring revealed that INTI is a unifactor test and shows all the subtests do load quite highly on this factor (Table 13).

Table 13

Factor Matrix showing the loadings of subtests on a single factor

Subtests	Factor 1
Series	.85
Matrices	.88
Analogies	.89
Odd-one-out	.81
Similarities	.83

The results of factor analysis presented in Table 13A and 13B establish the factorial validity of the test and clearly support our assumption that INTI is a measure of fluid intelligence.

Convergent and Discriminant Validation

Another method adapted to determine construct validity of INTI was the convergent and discriminant validation approaches (Campbell and Fiske,1959). This investigation was based on the assumption that INTI will correlate highly with the variables with which it should theoretically correlate (convergent validity) and it should have a significantly low correlation with the variables with which it should not correlate theoretically (discriminant validity).

Convergent Validity

To find out the convergent validity, the Indigenous Nonverbal Test of Intelligence (INTI) was correlated with other measures of ability. The criterion measures used for this purpose were Intelligence Test Battery (ITB) which is used by Army Selection and Recruitment Centres for preliminary selection of candidates for commission in Army and an adapted version of Raven's Standard Progressive Matrices (RSPM) used by Inter Services Selection Board (ISSB) for commission in Armed forces. The Intelligence Test Battery comprises two tests; a Verbal Intelligence Test (VIT) containing items of verbal comprehension, vocabulary, and numerical reasoning and a Nonverbal Intelligence Test (NIT) which contains analogies.

The newly devised Indigenous Nonverbal Test of Intelligence (INTI) and Intelligence Test Battery were administered at different Army Selection and Recruitment Centres to a sample of 200 candidates on the same day. The INTI and RSPM were administered at Inter Services Selection Board (ISSB) to another sample of 200 subjects. The RSPM was given first as part of the selection procedure at ISSB and INTI was administered to same group after an interval of two days.

To determine the convergent validity of INTI, test scores were correlated with the available data obtained from Army Selection and Recruitment Centres and ISSB. The results are given in Table 14.

Table 14

Correlations between Indigenous Nonverbal Test of Intelligence (INTI) and other Standardized Intelligence Tests

INTI & Subtests	*VIT*	*NIT*	*ITB*	*RSPM*
Series	.65***	.73***	.71***	.77***
Matrices	.69***	.74***	.72***	.82***
Analogies	.70***	.77***	.74***	.75***
Odd-one-out	.66***	.71***	.69***	.74***
Similarities	.64***	.72***	.69***	.74***
Full Test	.72***	.81***	.79***	.83***

***$p < .001$

The results presented in Table 14 show highly significant correlations ($p<.001$) between INTI, its subtests and other intelligence tests (VIT, NIT, ITB and RSPM). These findings provide strong evidence of convergent validity of the test.

Discriminant Validity

In order to find out discriminant validity, the INTI was correlated with a test measuring a different construct. The criterion measure used for the purpose was Individual Obstacles Test (IOs) administered at ISSB as part of selection procedure for commission in Armed forces. This test is used to assess candidate's mental and physical strength, mind-body coordination, sense of judgment and proportion, agility, courage, planning ability and reasoning. The INTI and IOs Test were administered to a sample of 200 subjects on the same day.

Table 15

Correlation between Indigenous Nonverbal Test of Intelligence (INTI) and Individual Obstacles Test (IOs)

INTI & Subtests	*IOs*	*p*
Series	.030	n.s
Matrices	.036	n.s
Analogies	.025	n.s
Odd-one-out	.018	n.s
Similarities	.033	n.s
Full Test	.013	n.s

As shown in Table 15, the low correlations between INTI, its subtests and IOs demonstrate the discriminant validity of INTI.

Criterion-Related Validity

One of the criteria most frequently employed in validating intelligence tests is some index of *academic achievement.* The specific indices used as criterion measures include school grades, achievement test scores, promotion and graduation records, special honors and awards, and teacher's ratings for *intelligence.* In the present study, the criterion of academic achievement used was marks in the preceding Higher Secondary School Certificate (HSSC) examination. The candidates who had appeared for HSSC examination and were waiting for results, their marks in first year examination were taken as criterion.

Table16

Correlation between Indigenous Nonverbal Test of Intelligence (INTI) and HSSC Examination Marks (N=200)

INTI & Subtests	HSSC Marks
Series	.67***
Matrices	.68***
Analogies	.70***
Odd-one-out	.65***
Similarities	.71***
Full Test	.74***

*** *P < .001*

The results presented in Table 16 reveal significantly high correlation ($p < .001$) between INTI and the criterion i.e., marks in the last annual examination. These findings yield evidence of test validity.

Intercorrelations among the Subtests

In order to find out the internal consistency of the INTI, intercorrelations among five subtests and the full test were computed.

Table 17

Intercorrelations among the Subtests and Indigenous Nonverbal Test of Intelligence (INTI)

INTI & Subtests	I	II	III	IV	V
1. Series	1.00				
2. Matrices	.77***	1.00			
3. Analogies	.77***	.79***	1.00		
4.Odd-one-out	.69***	.70***	.73***	1.00	
5. Similarities	.69***	.74***	.73***	.70***	1.00
6. Full Test	.89***	.91***	.91***	.86***	.86***

*** *P< .001*

Results presented in Table 17 show significant correlations among all the subtests as well as the full test ($p < .001$). These findings demonstrate internal consistency of the test suggesting that all the subtests measure the same general ability and similar cognitive functions. An inspection of Table 17 reveals that correlation between each one of the subtests and the full test is much higher than the intercorrelations among the subtests.

Other Statistical Analyses

Following statistical analyses were carried out to explore possible differences in the performance of different groups on Indigenous Nonverbal Test of intelligence. The comparison was made in terms of:

1. Urban versus Rural residence
2. Medium of Instruction (Urdu versus English) used in educational institutions where the candidates had studied
3. Science versus Arts groups

Urban versus Rural Residence

To determine the significance of differences in the mean scores of urban and rural samples on Indigenous Nonverbal Test of intelligence and its subtests, t-test for independent groups was used.

Table 18

Differences between the Mean Scores of Urban and Rural Samples on Indigenous Nonverbal Test of Intelligence (INTI)

	Urban (n = 100)		Rural (n = 100)				
INTI & Subtests	*M*	*SD*	*M*	*SD*	*df*	*t*	*p*
Series	9.99	2.85	9.66	2.75	198	.83	.406
Matrices	9.58	3.23	9.53	2.94	198	.11	.909
Analogies	10.37	3.19	9.89	2.83	198	1.12	.262
Odd-one-out	8.93	2.52	8.40	2.63	198	1.45	.148
Similarities	7.86	2.28	7.59	2.08	198	.87	.385
Full Test	46.73	12.47	45.07	11.82	198	.97	.335

The data given in Table 18 indicate that the mean scores of the subjects from urban areas on full test and all its subtests are marginally higher than the mean scores of the subjects from rural areas. However, no significant differences exist in the mean scores of urban and rural samples on full test and its subtests.

Urdu versus English Medium of Instruction

To determine the significance of differences in the mean scores of subjects from Urdu and English medium institutions on Indigenous Nonverbal Test of Intelligence and its subtests, t-test for independent groups was used.

Table 19

Differences between the Mean Scores of subjects from Urdu and English medium Institutions on Indigenous Nonverbal Test of Intelligence (INTI)

	Urdu Medium (n = 112)		English Medium (n = 88)				
INTI & Subtests	*M*	*SD*	*M*	*SD*	*df*	*t*	*p*
Series	9.69	2.68	9.92	2.89	198	.59	.556
Matrices	9.31	2.98	9.74	3.16	198	.96	.338
Analogies	9.92	2.82	10.29	3.17	198	.87	.386
Odd-one-out	8.44	2.55	8.83	2.61	198	1.07	.284
Similarities	7.42	2.11	7.96	2.22	198	1.75	.081
Full Test	44.79	11.78	46.76	12.40	198	1.14	.255

The results given in Table 19 show that the mean scores of subjects of English medium institutions on full test and all its subtests are slightly higher than the mean scores of Urdu medium institutions. However, all of these differences are insignificant.

Science Group versus Arts Group

To determine the significance of differences in the mean scores of subjects from Science group and Arts group on Indigenous Nonverbal Test of Intelligence and its subtests, t-test for independent groups was used.

Table 20

Differences between the Mean Scores of subjects from Science Group and Arts Group on Indigenous Nonverbal Test of Intelligence (INTI)

	Science Group (n = 122)		Arts Group (n = 78)				
INTI & Subtests	*M*	*SD*	*M*	*SD*	*df*	*t*	*p*
Series	10.52	2.78	8.93	2.58	198	4.16	.000
Matrices	10 .23	3.24	8.69	2.64	198	3.60	.000
Analogies	10.97	2.95	9.05	2.76	198	4.68	.000
Odd-one-out	9.15	2.64	8.04	2.39	198	3.06	.002
Similarities	8.35	2.12	6.92	2.07	198	4.86	.000
Full Test	49.24	12.29	41.64	10.58	198	4.61	.000

The results presented in Table 20 indicate significant differences between the mean scores of Science group and Arts group. These findings indicate that Science group is better than the Arts group on measured abilities.

STANDARDIZATION OF THE TEST

STANDARDIZATION OF THE TEST

One of the advantages of psychological tests in comparison with other forms of measurement is that tests are standardized. Standardization implies *Uniformity of procedure* in administering and scoring the test. To achieve uniformity of testing conditions, the author himself administered the newly developed test (INTI) to all the samples used for test development as well as the standardization sample. Furthermore, detailed instructions for administering the new test were prepared to secure uniformity of testing conditions. It was empirically established that the instructions are simple, clear and unambiguous. To attain uniformity in scoring, hand scoring keys were prepared for the test.

Another important step in the standardization of a test is the establishment of *norms.*

Development of Norms

Norms refer to the normative distribution, which shows the frequencies of persons obtaining specific scores on the test. Norms are developed by administering the test to a standardization sample, representative of the population for whom the test is to be designed.

Nearly all standardized tests now provide some form of within–group norms. With such norms, the individual's performance on a particular test is evaluated in terms of the performance of the most nearly comparable standardization group

(Anastasi & Urbina, 1997). The following types of norms for the Indigenous Nonverbal Test of Intelligence were developed:

1. Percentiles
2. Standard scores (T scores)
3. The Deviation IQ

Standardization Sample

The standardization sample was representative of the population for which the test is intended to be used. A stratified random sample of 1000 boys was selected from all the candidates who appeared for preliminary examination for commission in the Army at different Army Selection and Recruitment Centres all over the country. This sample comprised of candidates from the rural and urban areas of the four provinces of the country. The subjects comprised those candidates who had either passed HSSC or had appeared in the examination and were waiting for the results. They belonged to different socioeconomic and educational backgrounds representing all the four provinces of Pakistan. Age range of the sample was17 to 22 years (Annexure-C). Table 21 and 22 contain detailed information about the standardization sample of the study.

Table 21

Standardization Sample by Urban-Rural Residence and Provinces

Residence	Provinces				Total
	Punjab	Sind	NWFP	Baluchistan	
Urban	225	100	100	75	500
Rural	225	100	100	75	500
Total	450	200	200	150	1000

Table 22

Standardization Sample by Urban-Rural Residence and Age

Residence	Age Groups			Total
	17-18	19-20	21-22	
Urban	212	244	44	500
Rural	181	246	73	500
Total	393	490	117	1000

Procedure

Before beginning the test, the subjects were explained with the help of few examples, how to select/identify the appropriate response and record it on the answer sheets. They were also told about the time limit at the start of each subtest. Table 23 indicates the distribution of scores of standardization sample on Indigenous Nonverbal Test of Intelligence (INTI).

Table 23

Distribution of scores of standardization sample on Indigenous Nonverbal Test of Intelligence (N = 1000)

Scores	Frequencies	Percentages
0-15	0	0
16-20	10	1
21-25	30	3
26-30	50	5
31-35	77	7.7
36-40	116	11.6
41-45	119	11.9
46.50	160	16
51-55	150	15
56-60	135	13.5
61-65	83	8.3
66-70	42	4.2
71-75	24	2.4
76-80	4	0.4
81-90	0	0

Table 23 shows the distribution of scores of standardization sample on INTI. There is a fairly normal distribution of scores in different ranges. Percentage of subjects scoring in the middle ranges is higher as compared to upper and lower ranges.

Percentiles

Percentiles are derived scores expressed in terms of the percentage of persons in the standardization sample whose scores fall below a given raw score. Percentile scores have several advantages. They can be readily understood even by untrained persons. They are universally applicable, can be used equally well with adults and children, and are suitable for any type of test.

The percentile scores corresponding to raw scores of the subjects on the Indigenous Nonverbal Test of Intelligence were computed for the standardization sample (Table 24). The highest and the lowest scores obtained by the subjects on the test were 80 and 16 respectively.

Table 24

Percentiles corresponding to Raw Scores of the Standardization Sample on Indigenous Nonverbal Test of Intelligence (N = 1000)

Raw Scores	Percentiles	Raw Scores	Percentiles
20.00	1	50.00	55
27.00	5	52.00	60
31.00	10	54.00	65
35.00	15	55.00	75
37.00	20	57.00	75
39.00	25	58.00	80
41.30	30	61.00	85
43.00	35	63.00	90
45.00	40	68.00	95
47.00	45	74.00	99
49.00	50		

The results presented in Table 24 indicate that a raw score of 49 on Indigenous Nonverbal Test of Intelligence (INTI) corresponds to 50th percentile and hence, represents average performance. The raw scores in the range of 39 to 47 (between 25th to 45th percentiles) show below average performance on the test and scores below 39 (below 25th percentile) as poor performance. Similarly, subjects obtaining scores from 50 to 55 (above 50th to 74th percentiles) and 56 and above (75th percentile and above) reveal above average and outstanding intellectual ability respectively.

Standard Scores

The most commonly used type of normative scores are the standard scores, which express the distance of the individual's score from the mean of the distribution in standard deviation unit. A standard score may be obtained by either linear or nonlinear transformation of the original raw scores. However, most of the standard scores are derived using linear transformation of the original raw scores. In such transformation, they retain the exact numerical relations of the original raw scores, as they are computed by subtracting a constant from each raw score and then dividing the result by another constant.

The most basic standard score is the z score, which assumes a normal distribution having a mean of 0 and a standard deviation of 1. Any score can be converted into z score, simply by subtracting it from the mean of the scores and dividing by the standard deviation of the scores. The main disadvantage of the z score is that they include both negative numbers and decimal points.

T Scores

In order to provide standard scores using positive integer values, several other commonly used standard scores have been developed. If a normalized standard score i.e., z score is multiplied by 10 and added to or subtracted from 50, it is converted into a T score, a type of score first proposed by McCall (1922). Table 25 shows the T scores corresponding to different raw scores on the Indigenous Nonverbal Test of Intelligence.

Table 25

T scores corresponding to Raw Scores of the Standardization Sample on Indigenous Nonverbal Test of Intelligence (N = 1000)

Raw Scores	T Scores	Raw Scores	T Scores
1	12	15	23
2-3	13	16	24
4	14	17	25
5	15	18-19	26
6	16	20	27
7-8	17	21	28
9	18	22	29
10	19	23-24	30
11	20	25	31
12	21	26	32
13-14	22	27	33

Continue...

Raw Scores	T Scores	Raw Scores	T Scores
28	34	50-51	52
29-30	35	52	53
31	36	53	54
32	37	54	55
33	38	55	56
34-35	39	56-57	57
36	40	58	58
37	41	59	59
38	42	60	60
39	43	61-62	61
40-41	44	63	62
42	45	64	63
43	46	65	64
44	47	66	65
45-46	48	67-68	66
47	49	69	67
48	50	70	68
49	51	71	69

Continue…

Raw Scores	T Scores	Raw Scores	T Scores
72-73	70	82	78
74	71	83-84	79
75	72	85	80
76	73	86	81
77-78	74	87	82
79	75	88-89	83
80	76	90	84
81	77		

The Deviation IQ

The Deviation IQ is a standard score with a mean of 100 and SD that approximates the SD of Stanford-Binet distribution. Like Wechsler Intelligence Scales, a mean of 100 and SD of 15 have been used for deriving Deviation IQ scores for the Indigenous Nonverbal Test of intelligence (Table 26).

Table 26

Deviation IQs corresponding to Raw Scores of the Standardization Sample on Indigenous Nonverbal Test of Intelligence (N = 1000)

Raw Scores	Deviation IQs	Raw Scores	Deviation IQs
1	43	13	57
2	44	14	58
3	45	15	60
4	46	16	61
5	47	17	62
6	49	18	63
7	50	19	65
8	51	20	66
9	52	21	67
10	54	22	68
11	55	23	69
12	56	24	71

Continue...

Raw Scores	Deviation IQs	Raw Scores	Deviation IQs
25	72	47	99
26	73	48	100
27	74	49	101
28	76	50	103
29	77	51	104
30	78	52	105
31	79	53	106
32	81	54	107
33	82	55	109
34	83	56	110
35	84	57	111
36	85	58	112
37	87	59	114
38	88	60	115
39	89	61	116
40	90	62	117
41	92	63	118
42	93	64	120
43	94	65	121
44	95	66	122
45	96	67	123
46	98	68	125

Continue...

Raw Scores	Deviation IQs	Raw Scores	Deviation IQs
69	126	80	139
70	127	81	140
71	128	82	142
72	129	83	143
73	131	84	144
74	132	85	145
75	133	86	147
76	134	87	148
77	136	88	149
78	137	89	150
79	138	90	151

Interpretation of IQs

In consonance with the theory that the only unambiguous way to define intelligence levels is by delimiting them statistically. The Deviation IQ scores are, therefore, classified on a 7 point scale ranging from 1 to 7. This method of conversion was developed by British Army and is presently utilized by different other countries (e.g., Pakistan, Bangladesh, India), where the intelligence levels of the recruits/officers are assessed on 7 point scale.

IQ scores as given in Table 26 have been converted into 7 levels by arranging the distribution in order of size i.e., from IQ 151 to 43 and then assigning levels (1 to 7) in accordance with the normal curve percentages reproduced in Table 27. Table 28 indicates the classification of IQs into 7 intelligence levels.

Table 27

Normal curve percentages used for conversion of Deviation IQs into different Intelligence Levels

Percentages	5	10	15	40	15	10	5
Intelligence Levels	1	2	3	4	5	6	7

Table 28

Classification of Deviation IQs into Intelligence Levels

IQs Intervals	Intelligence Levels	Remarks	Percentages
127 and above	1	Outstanding	5
119-126	2	Well Above Average	10
111-118	3	Above Average	15
97-110	4	Average	40
87-96	5	Low Average	15
80-86	6	Below Average	10
79 and below	7	Well Below Average	5

Raw Score Equivalents of Percentiles, T Scores and Deviation IQs

Table 29 gives raw score equivalents of Percentiles, T Scores and Deviation IQs to facilitate interpretation of T Scores and Deviation IQs in terms of Percentiles.

Table 29

Raw Score Equivalents of Percentiles, T Scores and Deviation IQs

Raw Scores	Percentiles	T Scores	Deviation IQs
20	1	27	66
27	5	33	74
31	10	36	79
35	15	39	84
37	20	41	87
39	**25**	**43**	**89**
41	30	44	92
43	35	46	94
45	40	48	96
47	45	49	99
49	**50**	**51**	**101**
50	55	52	103
52	60	53	105
54	65	55	107
55	70	56	109
56	**75**	**57**	**111**
58	80	58	112
61	85	61	116
63	90	62	118
68	95	66	125
74	**99**	**71**	**132**

DISCUSSION

DISCUSSION

The Indigenous Nonverbal Test of Intelligence (INTI) has been devised to assess general intelligence of Pakistani youth. An effort was made to minimize the influence of verbal fluency, language, specific social and school learned skills and knowledge, experiential, cultural, socioeconomical and educational background on test scores of the subjects by using nonverbal test content comprising of figurative material selected from our own cultural and folk heritage. The main objective was to develop a test on scientific lines with sound psychometric characteristics and national norms.

The Final Version of INTI

The final version of the Indigenous Nonverbal Test of Intelligence (INTI) comprises five subtests: *Series*, *Matrices*, *Analogies*, *Odd-one-out* and *Similarities*. The total number of items in the test is 90. The first three subtests consist of 20 items each and the last two 15 each. The total test taking time for the test is 45 minutes.

Discrimination Index

Item discrimination indices of Indigenous Nonverbal Test of Intelligence (INTI) presented in Table 5 indicate that out of 90 items, 6 (6.7%) have very high discrimination power (.71 to .80), 22 (24.4%) have high discrimination power (.61 to

.70), 50 items (55.5%) have medium discrimination power (.41 to .60) and 12 (13.3%) show low discrimination power (.30 to .40). The average discrimination power of the test is .53. The highest discrimination power yielded by an item is .80 (Item 9, subtest Similarities) and the lowest is .30 (Item 16, Series; item 18, Analogies).

Items having very high (.71 to .80) and high (.61 to .70) ranges of discrimination indices are considered very good and effective items being powerful enough to discriminate between high and low scorers on the test. In other words, these items are very effective in discriminating between high and low ability levels on the test. Items with discrimination power in the middle range (.41 to .60) can also be considered as having moderate ability to discriminate between high and low performers on the test. However, items with low power of discrimination (.30 to .40) do not appear very effective in discriminating between high and low intellectual performances and need further evaluation and improvement.

Item Difficulty Level

Table 5 shows that out of 90 items, 4 (4.4%) have very high difficulty level (.71 to .80) and 20 (22.2%) have high difficulty level (.61 to .70). Whereas, 45 items (50%) have level of difficulty in the middle range (.41 to .60), 16 (17.8%) show lower level of difficulty (.31 to .40) and 5 items (5.5%) has item difficulty of .30. The average difficulty level (p value) of the test is .51. The most difficult item in the test has a difficulty level (p value) of.30, which shows that the probability of correct answer to this item is very low. Whereas, the easiest item has a difficulty level (p value) of .73, which indicates that the probability of its correct answer is very high.

Items in the high (.61 to .70) and middle (.41 to .60) ranges of difficulty level are good items being neither very easy nor very difficult. However, items with p values between .71 to .80 can be considered as very easy and items having p values between .30 to .40 as very difficult ones.

An overall picture of the quality of items in the test in terms of their discrimination power and difficulty level indicates that items in the test are almost normally distributed in different ranges of discrimination and difficulty. Items with very high, high, and medium discrimination power and with medium and low difficulty can be called really good and effective items. Those items with high and medium discrimination power and high difficulty level are also acceptable due to reasonably good discrimination power. The remaining categories of items though qualifying the selection criteria, need further revision and improvement. By and large, most of the items have turned out to be good and effective.

The results of item discrimination analysis presented in Table 8 demonstrate significant differences in the performance of high and low scoring groups on all the 90 items of the test. The high scoring group has obtained significantly higher scores on all the items. This analysis shows that all the items in the test are quite effective in discriminating high and low scorers and are capable of identifying differences in intellectual performance.

Comparison of Mean Percentages of correct responses of the Subtests

The comparison of mean percentages of correct responses to items comprising different subtests of INTI (Table 9) shows that *Similarities* has the lowest mean percentage of correct responses (47.7), while *Odd-one-out* has the highest mean percentage of correct responses (57. 7). These findings suggest that *Similarities* is the most difficult test whereas, *Odd-one-out* is the easiest one. These results may also be interpreted in terms of the familiarity of the content. Items comprising *Odd-one-out* may be depicting/presenting more familiar situations that has facilitated the subjects to deduce the underlying relationships.

Reliability of INTI

Three methods were used to establish reliability of the new test viz., KR-20; split-half and test-retest. The reliability indices estimated by all the three methods are quite high (.85 and above).

The procedures employed to find out different coefficients of reliability of the INTI are consistent with the large number of previous studies in which such procedures have been used. In a recent study, Hogan, Benjamin and Brezinski (2000) examined the frequency of use of various methods of reliability coefficients for a systematically drawn sample of 696 tests appearing in the APA-published Directory of Unpublished Experimental Mental Measures, Volume 7 (Goldman, Mitchel, & Egleson, 1997). After Alpha, the most frequently used reliability coefficients were Test-retest, Split-half and KR-20. Majority of the reliability coefficients recorded were in the range of .75 to .95, which also provides support to the reliability coefficients of the INTI.

Reliability is a property of the scores on the test for a particular population of examinees, not the test (Feldt & Brennan, 1989; Joint Committee on Standards for Educational Evaluation, 1994; Thompson & Vacha-Haase, 2000). Reliability of test changes with the changes in either (a) sample composition or (b) score variability

(Crocker & Algina, 1986). Therefore, it is always recommended that potential users of the test must compare the composition and variability of their sample with any previously reported reliability coefficient.

Validity of INTI

The data presented in Tables 12 to 16 concerning the validity of INTI obtained from various procedures establish the test as a valid measure of general intelligence. Different criteria of validation used in the present study included grade and age differentiation, correlation of INTI with other measures of general ability and marks in HSSC examination.

As hypothesized the mean scores of the subjects of the three different school grades and age groups on the subtests and the full test except *Similarities* show progressive increase with the advancing grade and age. Since, it is expected that children in different school grades will show variability in performance on ability tests and abilities are to be increased with age during childhood and adolescence, therefore, test scores are also expected to show such an increase if the test is valid. The progressive increase in the mean scores with different school grades and advancing age provides support for the validity of the test.

The difference in the mean scores of three different grades and age groups on *Similarities* has been found insignificant. Although the mean score of higher age group (grade12) is higher than the middle age group (grade11) and lower age group (grade10). Similarly the mean score of middle age group is also higher than the mean score of the lower age group.

The highly significant coefficients of correlation between INTI and other measures of intelligence ($p < .001$) as shown in Table 14 are indicative of convergent validity of the test. The other measures of intelligence used in the present study included Verbal Intelligence Test (VIIT), Nonverbal Intelligence Test (NIT), Intelligence Test Battery (VIT) and adapted version of Raven's Standard Progressive Matrices (RSPM). The positive and highly significant coefficients of correlations between the performances of the subjects on all the subtests of INTI and other measures of intelligence ($p < .001$) demonstrate the validity of the subtests as well. This also reflects that the performances on the subtests relate to the same underlying construct of intelligence. The high positive correlation between the scores of subjects on INTI and other measures of intelligence shows that the results of one test can be used to supplement or predict the performance of the subjects on the others. Correlations between INTI scores and other measures of intellectual abilities also suggest that the underlying processes may be general in nature, rather than specific to this particular measure. However, the criterion measures used for the validation of INTI like Nonverbal Test of Intelligence (NIT) and adapted version of Raven's Standard Progressive Matrices (RSPM) are limited in measuring general abilities simply because of insufficient variety of items. NIT includes only *Analogies* items and RSPM only *Matrices*. Therefore, inevitably these tests are likely to favour those subjects high on the factors specific to items of these tests and biased against those who are low on these factors. The newly developed Indigenous Nonverbal Test of Intelligence (INTI) consists of five different types of items involving different perceptual tasks, viz., *Series*, *Matrices*, *Analogies*, *Odd-one-out* and *Similarities.* It will provide an opportunity to measure different aspects of one's intellectual functioning. The items included in NIT and adapted version of RSPM are borrowed from Foreign tests as compared to INTI which includes items developed indigenously from figurative material of our own culture and the tasks involved in these items are equally valid for all segments of population. Furthermore, INTI is quite economical in terms of time required for administration as well as scoring.

The methods adopted for correlating INTI with other measures of intelligence are consistent with similar earlier studies. Wechsler (1955) correlated WAIS with Stanford-Binet and reported high correlation between Verbal, Performance and Full Scale IQs, and Stanford-Binet IQs (i.e., .86, .69 and .85 respectively).

Henderson (1964) studied more than 1000 Hong Kong primary school children and found a correlation of .51 between Raven's Matrices and scores on Form B of Cattell's Culture Fair Intelligence Test. Domino ((1964) adopted a similar procedure and reported a correlation of .56 and .55 between Cattell's Culture Fair Intelligence Test and the non-language form of California Test of Mental Maturity and GOH Schaldt Figures Test respectively. Similarly, Downing, Edgar, Harris, Kornberg and Storen (1965) observed high coefficients of correlation between Cattell's Culture Fair Intelligence Test and other measures of intelligence. For example, the researchers found a correlation of .49 with Otis Beta, .62 with WISC verbal, .63 with WISC performance and .72 with full scale WISC IQ and .56 with score on Metropolitan reading test.

Prewett (1992) examined the relationship between Kaufman Brief Intelligence Test (K-BIT) and WISC-R on a sample of 7-16 years old students. Both the tests were administered simultaneously. The K-BIT IQ composite correlated significantly with all the WISC-R full scale IQ scores.

The procedures adopted for determining the convergent validity of INTI are in line with earlier studies. The validity coefficients of INTI are almost of the same magnitude as found for other well known or extensively used standardized nonverbal tests of intelligence.

Table 16 shows highly significant correlation between INTI, its subtests and examination marks ($p < .001$). As a criterion measure, the examination marks are mostly regarded as the measure of performance and ability. Thus high correlation between INTI and examination marks establishes the criterion validity of INTI as a measure of intelligence. These findings suggest that INTI may be used to predict future performance of students in the examination. The students scoring high on INTI may score high in their college examination.

The significant correlations of INTI and its subtests with the examination marks are consistent with the results of past studies where nonverbal intelligence tests were correlated with school/college grades and other achievement measures. Saigh (1981) studied the validity of Lorge Thorndike Nonverbal Battery as a predictor of academic achievement and reported similar results. Students' final grades in English, Mathematics, Science, Social Sciences and their final GPA were moderately correlated with nonverbal IQ.

Prewett and Giannuli (1991) reported that the Verbal, Performance and Full Scale of WISC-R given to 66 students of ages 6 years 6 months to 11 years 11 months, who were referred to school psychologist and Stanford-Binet Intelligence Scale score of 48 referred students in the ages between 6 to 9 years, correlated significantly with the reading subtest of the Woodcock-Johnson Psycho-Educational Battery and Peabody's Achievement Test.

Morgan and Whorton (1991) determined criterion validity of WISC-R by a similar method as adopted for finding criterion validity of INTI. They administered WISC-R and Diagnostic Achievement Battery (DAB) to 12 native American and 13 white students aged 6 to 15 years, who were referred for psychological testing by the class teacher. The coefficients of correlation between DAB and WISC-R Verbal, Performance and Full scale IQ were found to be .78 and.69 respectively.

School marks have been used as a validity criterion in some other studies conducted in our culture. Israr and Abbas (1990) developed a Test of Intellectual Development for Pakistani Preschool Children and reported moderate to high correlations (.31 to .71)) between scores on this test and school marks for different samples. Syed (1993) found a correlation of .71 between the scores on Nonverbal Test of intelligence for Pakistani Urban Primary School Children and the school marks. Similarly, Gardezi (1994) devised a Nonverbal Intelligence Test for Adolescents and reported high correlation (.71) between the scores of the subjects on the test and their examination marks.

The results of the aforementioned studies indicate that the validity coefficients of INTI and its subtests are almost of the same magnitude as found for other nonverbal tests of intelligence. The similarity in the validity indices of INTI and other well established standardized nonverbal tests of intelligence yield evidence to support INTI as a valid measure of nonverbal intelligence.

Intercorrelations of the Subtests

Intercorrelation among the subtests and correlation between each subtest and the full test as shown in Table 17 indicate that all the five subtests i.e., *Series*, *Matrices*, *Analogies*, *Odd-one-out* and *Similarities* are highly correlated with each other as well as with the full test. The range of correlation among the subtests is .69 to .79 and correlation between each subtest and the full test is from .86 to .91. All the correlations are highly significant ($p < .001$) and reflect that all the subtest measure similar intellectual functions. The high correlations between each subtest and the full test also provide empirical evidence of the consistency of the test.

Urban versus Rural Residence

As shown in Table 18 there is no significant difference in the mean scores of urban and rural subjects on INTI and its subtests. The results are in accordance with our expectations and support main assumption of the test. The test has minimized the influence of experiential, cultural, socioeconomical and educational background of urban sample as compared to rural sample. The insignificant difference in the performance of urban and rural subjects reflects that the tasks involved in the items of INTI are equally valid for urban and rural population of adolescents.

The past research concerning assessment of nonverbal intelligence has reported significant differences in the intellectual performances of urban and rural subjects (Ansari & Iftikhar, 1984; Kauser, 1998). However, these studies were conducted on children of lower grades. As the children advance to higher grades and are exposed to wider educational, social and cultural experiences at higher schools/colleges, the urban and rural differences in their intellectual performances are minimized. Electronic media and other sources of information provide intellectual stimulation. Pakistani youth of today, whether urban or rural are exposed to a wider range of environmental stimuli than their ancestors are, and that has also improved the level and breadth of knowledge and general ability of rural youth.

Urdu versus English Medium of Instructions

The comparison of performance of Urdu and English medium students on INTI and its subtests (Table 19) does not reveal any significant difference. These results are in the expected direction and support the main objective of the study. The test, as in the case of urban and rural subjects has also minimized the advantage of experiential, socioeconomical, cultural and educational background of English medium students over Urdu medium students. In other words, it can be claimed that items included in the test are loaded on 'g' factor.

Generally in Pakistan, there is a better social and educational environment in English medium institutions as compared to Urdu medium, which favours child's mental growth. The students from English medium institutions have a definite edge over the students of Urdu medium institutions in areas like verbal fluency, general knowledge, and specific social and school learned skills, but as INTI is a nonverbal test of intelligence with figurative test contents, the expected advantage of the test being equally valid for persons varying in social and educational experiences has been established.

Science versus Arts groups

The comparison of performance of students of Science and Arts groups (Table 20) shows significant differences in their scores on INTI and its subtests. The mean scores of the Science group are higher than the mean scores of Arts group on all the five subtests as well as the full test. These findings reflect that items of the test are closer to the experiences of the Science students. The test contents, specifically geometrical figures, seem closer to different abstract concepts used by Science students, which might have facilitated their performance on the test. Another possible reason of this difference in the mean scores of two groups may be the fact that in Pakistan, it is mostly the brilliant students who study Science subjects and always secure higher positions in External Examinations (Board Examinations).

Conclusions

Taken together the results of different analyses, it can be concluded that INTI withstands the psychometric scrutiny and appears to be a comprehensive and valid measure of intelligence. It can be used for the assessment, comparison and prediction of intellectual potential of Pakistani youth in schools and colleges, personnel selection, and other research and applied settings.

In Pakistan, although many research studies have been conducted on adaptation and validation of intelligence tests and related issues, yet there is little genuine work found on the development of indigenous nonverbal intelligence tests, specifically for adolescents and adults. Some efforts have been made in the last few decades to adapt and validate a few famous nonverbal intelligence tests like Raven's Progressive Matrices and Cattell's Culture Fair Intelligence Test on different populations and against different criteria. However, most of these adaptation/validation studies were conducted on small samples representing specific groups.

Overall, the present research can be considered a first and pioneering effort towards the development of an indigenous nonverbal test of intelligence based on scientific psychometric procedures. However, as the standardization sample comprised of only those boys who were candidates for commission into Pakistan Army, the test in its present form is not recommended to be used for all adolescents of Pakistan till the development of the national norms.

Recommendations for Further Research

1. To increase the validity of INTI, items with relatively low discriminatory power may be modified or new items may be added. Similarly item choices, which failed to attract considerable number of respondents, may also be improved or replaced.

2. The target population for which INTI has been developed consisted of candidates who appeared at different Army Selection and Recruitment Centres for regular commission in Army. In Pakistan only males are eligible for commission in regular Army, though females are also inducted in some specialized branches. Therefore, INTI was standardized only on male sample. Similar test must be devised and standardized for female population.

3. To provide further evidence of construct validity, the INTI may be correlated with other standardized measures of intelligence like Raven's Advanced Progressive Matrices, Cattell's Culture Fair Intelligence Test, WAIS, Otis Lennon Mental Ability Test etc.

4. For further exploration concerning performance of Science group, a comparison of pre-medical and pre-engineering students on INTI may be carried out.

5. Age norms may be developed by administering INTI to different age groups.

6. To establish national norms, INTI may be administered on a stratified sample including about 10,000 subjects.

REFERENCES

REFERENCES

Ahmed, I. (1985). *Development of Urdu version of California Psychological Inventory in Pakistan.* National Institute of Psychology, Quaid-i-Azam University, Islamabad, Pakistan.

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

Aiken, L. R. (1991). *Psychological testing and measurement* (7th ed.). Boston: Allyn & Bacon.

Aiken, L. R. (1997). *Psychological testing and assessment* (9th ed.). Boston: Allyn & Bacon.

Ain, Q. (1985). A validation study of Cattell's Culture Fair Intelligence Test on Pakistani Children. In Z. A. Ansari, P. N. Tariq, R. M. Zaman, S. Chowdhri, & S. Watanabe (Eds.), *Proceedings of the fifth conference of Pakistan Psychological Assoication*, Islamabad, Pakistan.

Allinger, G. M. (1988). Do zero correlations really exits among measures of different intellectual abilities? *Educational and Psychological Measurement, 48,* 275-280.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1985). *Standards for educational and psychological testing.* Washington, DC: American Psychological Association.

Anastasi, A. (1983). What do intelligence tests measure? In S. B. Anderson & J. S. Helmick (Eds.), *On educational testing* (pp. 5-28). San Francisco: Jossey-Bass.

Anastasi, A. (1986). Intelligence as a quality of behaviour. In R. J. Sternberg & D. K. Detterman (Eds.), *What is intelligence? Contemporary viewpoints on its nature and definitions* (pp. 19-21). Norwood, NJ: Ablex.

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Andreasen, N. C., Flaum, M., Swayze, V., II, O'Leary, D. S., Alliger, R., Cohen, G., Ehrhardt, J., & Yuhn, W. T. C. (1993). Intelligence and brain structure in normal individuals. *American Journal of Psychiatry, 150*, 130-134.

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

Ansari, Z. A., & Iftikhar, M. (1984). Validity of Raven's Standard Progressive Matrices for urban and rural school children in Pakistan. Part1: Basic facts. *Psychology Quarterly, 19*, 14-27.

Ansari, Z. A., Tariq, P. N., & Iftikhar, M. (1990). *Educational Ability Test Level 5: Development and validation.* National Institute of Psychology, Quaid-i-Azam University, Islamabad, Pakistan.

Arlin, P. K. (1984). Adolescent and adult thought: A structural interpretation. In M. I., Commons, F. A. Richadrs, & C. Armon (Ed.), *Beyond formal operations: Late adolescent and adult cognitive development* (pp.239-252). New York: Praeger.

Armstrong, T. (1994). *Multiple intelligences in the classroom.* Virginia: Association for supervision and curriculum development.

Arnau, R., & Thompson, B. (2000). Second order confirmatory factor analysis of the WAIS-III. *Assessment 2000, 7*(3), 237-246.

Arvey, R. D. (1979). *Fairness in selecting employees.* Reading, M A: Addison-Wesley.

Aziz, S. (1997). *Adaptation of Columbia Mental Maturity Scale in Pakistan*. National Institute of Psychology, Quaid-i-Azam University, Islamabad, Pakistan.

Bachelor, P. A. (1989). Maximum likelihood confirmatory factor-analytic investigation of factors with Guilford's structure of intellect model. *Journal of Applied Psychology, 74,* 797-804.

Bailiargeon, R. (1991, April). *Infants' reasoning about collision events.* Paper presented at the Society for Research in Child Development, Seattle.

Baron, R. A. (1999). *Psychology* (3rd ed.). New Delhi: Prentice-hall of India.

Bayley, N. (1970). Development of mental abilities. In P. Mussen (Ed.), *Carmichoel's Manual of Child Psychology (*pp.1163-1209). New York: Wiley.

Berry, J. W., & Dasen, P. R. (1974). *Culture and cognition: Readings in cross cultural psychology*. London: Methaen.

Berry, J. W., Poortinga, Y. A., Segall, M. H., & Dasen, P. R. (1992). *Cross-cultural psychology: Research and applications*. New York: Cambridge University Press.

Bouchard, T. J., Jr., Lykken, D. T., McGue, M., Segan, N. L., & Tellegen, A. (1990). Sources of human psychological differences: The Minnesota study of twins reared apart. *Science*, *250*, 223-228.

Bouchard, T. J., Jr., & Segal, N. L. (1985). Environment and IQ. In B. B. Wolman (Ed.), *Handbook of intelligence: Theories, measurements, and applications*. New York: Wiley.

Boykin, A. W. (1994). Harvesting talent and culture: African-American children and educational reforms. In R. Rossi (Ed.), *Schools and students at risks.* New York: Teachers College Press.

Bracken, B. A., & McCallum, R. S. (1998). *Universal Intelligence Test examiner's manual.* IL: Riverside. + Retrieved January 6, 2001, from ericae on line database on the world wide web: ericae. com.

Brislin, R. W. (1990). *Applied cross-cultural psychology*. London: Sage Publications.

Brislin, R. W. (1993). *Understanding culture's influence on behavior*. Orlando, FL: Harcourt Brace Jovanovich.

Brody, N. (1992). *Intelligence* (2nd ed.). New York: Basic Books.

Brown, A. L. (1978). Knowing when, where and how to remember: A problem of Meta cognition. In R. Glaser (Ed.), *Advances in Instructional Psychology*. Vol I. Hillsdale, NJ: Erlbaum.

Brown, L., Sherbenou, R. J., & Johnsen, S. K. (1990). *Test of Nonverbal Intelligence-* (2nd ed.). Austin, TX: PRO-ED. + Retrieved January 6, 2001, from ericae on line database on the world wide web: ericae. com.

Buros, O. K. (1959). *The Fifth Mental Measurements Year Books*. New Jersey: The Gryphon Press.

Buros, O. K. (1965). *The Sixth Mental Measurements Year Books*. New Jersey: The Gryphon Press.

Burt, C. (1954). Differentiation of intellectual ability. *British Journal of Educational Psychology*, *24*, 76-90.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the Multitrait- Multimethod Matrix. *Psychological Bulletin*, *56*, 81-105.

Carlson, N. R. (1988). *Discovering psychology*. Boston: Allyn & Bacon.

Carpenter, P.A., Just, M. A., & Shell, p. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices test. *Psychological Review*, *97*, 404 431.

Carroll, J. B. (1976). Psychometric tests as cognitive tasks: A new structure intellect. In L. B. Resnick (Ed.), *The nature of intelligence*. Hillsdale, NJ: Erlbaum.

Carroll, J. B. (1980). *Individual differences in psychometric and experimental cognitive tasks*. (Final Report). Chapel Hill, NC: L. L. Thurstone. Psychometric Laboratory, University of North Carolina.

Carroll, J. B. (1981). Ability and task difficulty in cognitive psychology. *Educational Research*, *10*, 11-21.

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

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

Cattell, R. B. (1967). The theory of fluid and crystallized intelligence. *British Journal of Educational Psychology*, *37*, 209-224.

Cattell, R. B. (1971). *Abilities; Their structure, growth and action*. New York: Houghton.

Cattell, R., & Horn, J. (1978). A check on the theory of fluid and crystallized intelligence with description of new subtest designs. *Journal of Educational Measurement*, *15*, 139-164.

Cattell, R. B. (1987). *Intelligence*. New York: Elsevier.

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology*, *27*, 703-722.

Chance, P., & Fischman, J. (1987, May). The magic of childhood. *Psychology Today*, 48-58.

Cleary, T., Humphreys, L., Kendrick, S., & Wesman, A. (1975). Educational use of tests with disadvantages students. *American Psychologist*, *30*, 15-41.

Cohen, R. J., Montague, P., Nathanson, L. S., & Swerdlik, M. E. (1988). *Psychological testing: An introduction to test & measurement*. Mountain View, CA: Mayfield.

Cole, M., & Cole, S. R. (1989). *The development of children.* New York: Scientific American.

Commons, M. L., Sinnott, J., Richards, F. A., & Armion, C. (Eds.). (1989). *Beyond formal operations: Comparison and applications of adolescent and adult developmental models*. New York: Praeger.

Conoley, J. C., & Kramer, J. J., & Mitchell, J. V. (1988). *The supplement to the ninth mental measurements year books*. Lincotin: The University of Nebraska Press.

Conoley, J. C., & Kramer, J. J. (1989). *The tenth mental measurements year books.* Lincotin: The University of Nebraska Press.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York:Holt, Rinehart & Winston.

Cronbach, L. J. (1975). Five decades of public controversy over psychological testing. *American Psychologist*, *30*, 1-14.

Cronbach, L. J. (1984). *Essentials of psychological testing.* (4th ed.). New York: Harper & Row.

Cronbach, L. J., & Snow, R. E. (1977). *Attitudes and instructional methods: A handbook for research and interactions*. New York: Irvington.

Dai, x. y., Ryan, J. J., Paolo, A. M., & Harrington, R. G. (1990). Factor analysis of the Mainland Chinese Version of the WAIS-RC in a brain damaged sample. *International Journal of Neuroscience, 55*(2-4), 107-111.

Davie, R., Butter, N., & Goldstein, H. (1972). *From birth to seven: A report of the National Child Development Study*. London: Longman and the National Children's Bureau.

Deary, I. J. (1995). Auditory inspection time and intelligence: What is the direction of causation? *Developmental Psychology*, *31*, 327-350.

Deary, I. J., & Stough, C. (1996). Intelligence and inspection time. *American Psychologist*, *51*, 599-608.

DeFries, J. C., & Plomin, R. (1978). Behavioral genetics. *Annual Review of Psychology, 29*, 473-515.

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

Domino, G. (1964). Comparison of the D 48, Cattell Culture Fair and Army Beta Test in a sample of college males. *Journal of Consulting Psychology*, *28*, 468-469.

Douglas, J. W. B. (1969). *The home and the school*. London: MacGibbon & Kee.

Douglas, J. W., Ross, J. M., & Simpson, H. R. (1968). *All our future.* London: Peter Davies.

Downing, G. L., Edgar, R. W., Harris, A. J., Kornberg, L., & Storen, H. F. (1965). *The preparation of teachers for schools in culturally deprived neighborhoods:* Cooperative Research Project 935. Queens Collage of the City University of New York.

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

Dunn, L. M., & Dunn, L. M (1981). *Peabody Picture Vocabulary Test.* Revised: Manual for Forms L and M. Circle Pines, MN: American Guidance Service.

Enns, R. A., & Reddon, J. R. (1998). The factor structure of the Wechsler Adult Intelligence Scale-Revised: One or two but not three factors. *Journal of Clinical Psychology, 54* (4), 447-459.

Eysenck, H. J., & Kamin, L. (1981). *Intelligence: The battle for the mind.* London: Pan Books. .

Feldt, L. S., & Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 105-146). Washington, DC: American Council on Education/Oryx.

Flavell, J. H. (1985). *Cognitive development* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Flavell, J. H. (1993). Young children's understanding of thinking and consciousness. *Current Directions in Psychological Science, 2*, 40-43.

Flynn, J. R. (1994). IQ gains overtime. In R. J. Sternberg (Ed.), *Encyclopedia of human intelligence* (pp. 617-623). New York: Macmillan.

Frensch, P. A., & Sternberg, R. J. (1990). Intelligence and cognition. In M. W. Eysenck (Ed.), *Cognitive psychology: An international review,* (pp. 57-103). New York: Wiley.

Fry, A. F., & Hale, S. (1996). Processing speed, working memory and fluid intelligence. *Psychological Science*, 7, 237-241.

Gardezi, A. H. (1994). *Development of a Non verbal Intelligence Test for adolescents.* Unpublished M. Phil. dissertation. National Institute of Psychology, Quaid-I-Azam University, Islamabad, Pakistan.

Gardner, H. (1983). *Frames of mind.* New York: Basic Books.

Gardner, H. (1986). From testing intelligence to assessing competencies. A pluralistic view of intellect. *Roeper Review*, *8*(3), 147-150

Gardner, H. (1993). *Multiple intelligences: The theory in practice*. New York: Basic Books.

Glazer, S. (1993). Intelligence testing. *CQ Researcher*, *3*, 651-659.

Goldman, B. A., Mitchel, D. F., & Egelson, P. E. (1997). *Directory of unpublished experimental measures* (Vol. 7). Washington, DC: American Psychological Assoication.

Guilford, J. P. (1988). Some changes in Structure-of-Intellect model. *Educational and Psychological Measurement, 48,* 1-4.

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

Guion, R. M. (1965). *Personnel testing*. New York: McGraw-Hill.

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

Haier, R. J. (1996). Cerebral glucose metabolism and Intelligence. In P. A. Vernon (Ed.), *Biological approach to the study of human intelligence*. Norwood, NJ: Ablex.

Hammill, D. D. (1996). *Comprehensive Test of Nonverbal Intelligence*. (Retrieved January 6, 2001, from ericae on line database on the world wide web: ericae. com).

Hartigan, J. A., & Widgor, A. K. (1989). *Fairness in employment testing: Validity generalization, minority issues, and the General Aptitude Test Battery.* Washington, DC: National Academy Press.

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

Heim, A. W. (1975). *Psychological testing.* London: Oxford University Press.

Henderson, N. K. (1964). *The attainment and ability of Hong Kong primary IV pupils* (A Godman ed.). Hong Kong: Hong Kong Council for Educational Research.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American society.* New York: Free Press.

Hogan, T. P., Benjamin, A., & Brezinski, K. L. (2000). Reliability methods: A note on the frequency of use of various types. *Educational and Psychological Measurement*, *60*, 523-531.

Horn. J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligence. *Journal of Educational Psychology*, *57*, 253-270.

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica*, *26*, 107-129.

Horn, J. L. (1978). *Individual differences in cognitive development*: Final report. Betheoda, Maryland: National Science Foundation.

Horn, J. L. (1981). *Age difference in human abilities*: Final report. Bethesda, Maryland: National Institute on Aging.

Huffman, K., Vernoy, M., & Vernoy, J. (1994). *Psychology in action* (3rd ed.). New York: John Willey & Sons, Inc.

Humphreys, L. G. (1979). The construct of general intelligence. *Intelligence*, *3*, 105-120.

Hussan. 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.

Hussain, S. S. (1992). *Development of Group Verbal Intelligence Test in Urdu for high school students*. Unpublished M.Phil. dissertation. National Institute of Psychology, Quaid-i-Azam University, Islamabad, Pakistan.

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 Raven's Standard Progressive Matrices. *Pakistan Journal of Psychology*, *17* (3-4), 23-35.

Israr, M. (1985). Cognitive development of Pakistani primary school children: Salient findings of a research project. In Z. A. Ansari: P. N. Tariq: R. M. Zaman: S. Chowdhri and S. Watanabe (Eds), *New directions in Pakistani psychology: Proceedings of the fifth Conference of Pakistan Psychological Association* (pp. 67-75) Islamabad: PPA.

Israr, M. (1988). *A psychological interpretation of mathematics learning problems among secondary school students*. Unpublished M. Phil. dissertation. National Institute of Psychology, Quaid-i-Azam University, Islamabad, Pakistan.

Israr, M., & Abbas, I. (1990). *Test of Intellectual Development for Pakistani Pre-school Children: Development and validation.* National Institute of Psychology, Quiad-i-Azam University, Islamabad, Pakistan.

Jackson, G. D. (1975). On the report of the ad hoc committee on educational uses of tests with disadvantaged students. *American Psychologist, 30*, 88-93.

Jamal, S. (1964). *The validity of Raven's Colored Progressive Matrices Test and Otis Quick Scoring Mental Ability Test for Pakistani children.* Unpublished M.A. thesis. Institute of Education and Research: University of the Punjab, Lahore, Pakistan.

Jensen, A. R. (1969). How much can we boost IQ scholastic achievement? *Harvard Educational Review, 39*, 1-123.

Jensen, A. R. (1973). *Educability and group differences*. London: Methuen.

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

Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards: How to assess evaluation of educational programs* (2nd ed.). Thousand Oaks, CA: Sage.

Joseph, A. (1977). *Intelligence, IQ, and race-when, how, and why they became associated.* San Francisco: R & E Research Associates.

Kalyan, M. V. (1962). *Hd Km Non-Verbal Test, Forms A and B*. Princeton, NJ: Rosedale and Carter. + Retrieved January 6, 2001, from ericae on line database on the world wide web: ericae. com.

Katzell, R., & Dyer, F. (1977). Differential validity revived. *Journal of Applied Psychology, 62*, 137-145.

Kaufman, A. S., & Kaufman, N. L. (1983). *Kaufman Assessment Battery for Children: Administration and scoring manual.* Circle Pines, MN: American Guidance Service.

Kausar, R. (1998). *Validity study of Raven's Coloured Progressive Matrices (CPM) in Pakistan.* Unpublished M. Phil. Dissertation. National Institute of Psychology, Quaid-I-Azam University, Islamabad, Pakistan.

Khalid, H. (1986). *Differences between primary students of government and private schools in cognitive development.* Unpublished M. Phil. dissertation. National Institute of Psychology, Quaid-i-Azam University, Islamabad, Pakistan.

Khan, S., & Ahmad, I. (1985). Adaptation of Columbia Mental Maturity Scale in Pakistan. *Technical Report* (p.14). Islamabad: National Institute of Psychology.

Klemchuk, H. P., Bond, L. A., & Howell, D. C. (1990). Coherence and correlates of level 1 perspective taking in young children. *Merrill-Palmer Quarterly, 36*, 369-387.

Klineberg, O. (1928). An experimental study of speed and other in "racial" differences. *Archives of Psychology*, No. 93.

Kline, P. (1986). *A handbook of test construction: Introduction in psychometric design.* London: Methuen.

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

Knapp, R. R. (1960). The effects of time limits on the intelligence test performance of Mexican and American subjects. *Journal of Educational Psychology, 51*, 14-20.

Lerner, R. M. (1990). Plasticity, person-context relations and cognitive training in the aged years: A developmental contextual perspective. *Developmental Psychology, 26*, 911-915.

Linn, R. (1982). Ability testing: Individual differences, prediction, and differential prediction. In A. Wigdor and W. Garner (Eds.), *Ability testing: Uses, consequences, and controversies* (Part 11). Washington, D. C: National Academy Press.

Lubinski, D., & Benbow, C. P. (1995). An opportunity for empiricism, Review of Multiple intelligences: The theory and practice. *Contemporary Psychology, 40*, 935-940.

Lynn, R. (1987). The intelligence of 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 Spearmen's g and the neurospatial and verbal primaries. *Intelligence*, *12*, 423-433.

Mahmud, N. (1990). Development and Validation of a Educational Ability Test for Pakistani pre-school children. Unpublished M.Phil. dissertation. National Institute of Psychology, Quaid-i-Azam University, Islamabad, Pakistan.

Mahmood, Z. (1991). Intelligence, IQ and the Third World. *Pakistan Journal of Psychological Research*, *6, (*1-2*)*, 31-53.

Marks-Tarlow, T. (1996). *Creativity inside out: learning through multiple intelligences*. New York: Addison-Wesley.

Matarazzo, J. D. (1992). Psychological testing and assessment in the 21st century. *American Psychologist*, *47*, 1007-1018.

McCall, W. A. (1922). *How to measure in education.* New York: Macmillan.

McNemar, Q. (1942). *The revision of the Standford-Binet Scale*. Boston: Houghton Miffin.

Meltzoff, A. N., & Moore, M. K. (1989). Imitation in newborn infants: Exploring the ranges of gestures imitated and the underlying mechanisms. *Developmental Psychology*, *25*, 954-962.

Messick, S. (1992). Multiple intelligences or Multilevel intelligences? Selective emphasis on distinctive properties of hierarchy: On Gardner's Frames of mind and Sternberg's Beyond IQ in the context of theory and research on the structure of human abilities. *Psychological Inquiry, 3* (4), 365-384.

Morgan, R. L., & Whorton, J. E. (1991). Comparison of the Diagnostic Achievement Battery and WISC-R: A pilot study, *Psychological Reports*, *69*, 636-638.

Murphy, K. R., & Davidshofer, C. O. (1998). *Psychological testing: Principles and application* (4th ed.). New Jersey: Prentice-Hall.

Naglieri, J. A. (1985). *Matrix Analogies Test-Expanded Form: Examiner's manual.* San Antonio, TX: Psychological Corporation.

Naglieri, J. A., & Das, J. P. (1990). Planning, attention, simultaneous and successive (PASS) cognitive processes. *Journal of Psychoeducational Assessment, 8*, 303-337.

Naglieri, J. A. (1997). Knowns and unknowns, hits and misses. *American Psychologist*, *52*, 75-76.

Naglieri, J. A., & Bardos, A. N. (1998). *General Ability Measure for Adults*. (Retrieved January 6, 2001, from ericae on line database on the world wide web: ericae. com).

Naglieri, J. A., & Prewett, P. N. (1999). Nonverbal intelligence measures: A selected review of instruments and their use. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Intelligence and achievement* (pp. 348-370). New York: Guilford.

Naheed, G. (1993). *Development of a Verbal Test of Intelligence for Pakistani Urban Primary School Children*. Unpublished M. Phil. dissertation. National Institute of Psychology, Quaid-i-Azam University, Islamabad, Pakistan.

Neisser, U., Boodoo, G., Bouchard, T. J. Jr., Boykin, A. W., Brody, N., Ceci, S.J., Halpern, D. E., Loehlin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*, 77-101.

Newell, A., & Simon, H. A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice- Hall.

Ogbu, J. U. (1994). Fromcultural differences to differences in cultural frame of reference. In P. M. Greenfield & R.R. Cocking (Eds.), *Cross-cultural roots of minority child development* (pp.365-391). Hillsdle, NJ: Erlbaum.

Otis, A., & Lennon, R. J. (1968). *Manual for administration.* New York: Psychological corporation.

Parmar, R. S. (1989). Cross-cultural transfer of Non-Verbal Intelligence Tests. *British Journal of Educational Psychology*, *59*, 379-388.

Pellegrino, J. W. (1985). Inductive reasoning ability. In R.J. Stenberg (Ed.). *Human abilities: An information processing approach*. San Francisco: Freeman.

Piaget, J. (1952). *The child conception of numbers.* London: C. Gallego and F. Hodgon.

Piaget, J. (1976). *Psychology of Intelligence*. Totowa, NJ: Little Field, Adams.

Pidgeon, D. A. (1970). *Expectation and pupil performance*. Windson: NFER.

Piedmont, R. L., Sokolove, R. L., & Fleming, M. Z. (1992). Analysis of various WAIS-R factor structures in a psychiatric sample. *Journal of Clinical Psychology*, *48*(5), 658-66.

Plomin, R. (1989). Environment and genes: Determinants of behaviour. *American Psychologist*, *44*, 105-111.

Prewett, P. N. (1992). The relationship between the Kaufman Brief Intelligence Test (K-BIT) and the WISC-R with referred students. *Psychology in the Schools*, *29*, 25-27.

Prewett, P.N., & Giannuli, M. M. (1991). Correlation of WISC-R, Stanford-Binet Intelligence Scale and the reading subtests of the popular achievement tests. *Psychological Report*, *64*, 1232-34.

Pyle, D. W. (1979). *Intelligence: An introduction.* London: Routledge & Kegan Paul.

Raven, J. (1989). The Raven Progressive Matrices: A review of national norming studies and ethnic and socioeconomic variation within the United States. *Journal of Educational Measurement*, *26*, 1-16.

Raven, J. C. (1983). The Progressive Matrices and Mill Hill Vocabulary Scale in western societies. In S. H. Irvine & J. W. Barry. (Eds.), *Human assessment and cultural factors*. New York: Plenum.

Raven, J., Raven, J. C., & Court, J. H. (1995). *Manual for Raven's Progressive Matrices and Vocabulary Scales.* Oxford, England: Oxford Psychologists Press.

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

Rowe, H. A. H. (Ed). (1991). *Intelligence: Reconceptualization and measurement.* Hillsdale, NJ: Erlbaum.

Saigh, P. A. (1981). The validity of the Lorge Thorndike Nonverbal Battery as a predictor of the academic achievements of internal students. *Journal of Educational & Psychological Measurement, 42*, 1315-18.

Salovery, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality, 9*, 185-211.

Sattler, J. M. (1992). *Assessment of children* (3rd ed., revised). San Diego, CA: Author.

Schaie, K. W. (1983). *Schaie: Thurston Adult Mental Abilities Test: Manual.* Palo Alto, CA: Consulting Psychologists Press.

Schmitt, N., Mellon, P., & Bylenga, C. (1978). Sex differences in validity for academic and employment criteria and different types of predictions. *Journal of Applied Psychology, 63*, 145-150.

Schwarz, P. A., & Krug, R. E. (1972). *Ability testing in developing countries: A handbook of principles and techniques*. New York: Praeger.

Shah, A. A. (1994, April). *21st century: Needs and requirements in psychological measurement and evaluation.* Paper presented at Inter Services Selection Board Seminar, Kohat, Pakistan.

Sheikh, M. H., & Khan, K. H. R. (1982). *Zahanat Paima: Translation and adaptation of Otis Quick Scoring Mental Ability Test (Form C&D)*. Lahore: Department of Applied Psychology, University of the Punjab.

Snow, R. E. & Yalow, E. (1982). Education and intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 493-571). Cambridge: Cambridge University Press.

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

Spelke, E. S. (1988). Where perceiving ends and thinking begins: The apprehension of objects in infance. In A. Yonas (Ed.), *Minnesota symposia on child psychology* (Vol 20: Perceptual development in infancy). Hillsdale, NJ: Erlbaum.

SPSS, Inc. (1990). *SPSS Reference Guide*. Chicago: SPSS Inc.

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

Sternberg, R. J. (1980). Sketch of a componential sub-theory of human intelligence. *Behavioural and Brain Sciences, 3*, 573-614.

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

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

Sternberg, R. J. (1990). *Metaphors of mind: Conceptions of the nature of intelligence.* New York: Cambridge University Press.

Sternberg, R. J. (1994). *Encyclopedia of human intelligence*. New York: Macmillan.

Sternberg, R. J., & Detterman, D. K. (Eds.). (1986). *What is intelligence? Contemporary viewpoints on its nature and definition*. Norwood, NJ: Albex.

Sternberg, R. J., Wagner, R. K., William, W. M., & Horvath, J.A. (1995). Testing Common Sense. *American Psychologist, 50*, 912-927.

Stodard, G. D. (1943). *The meaning of intelligence*. New York: Macmillan.

Sugarman, S. (1987). *Piaget's construction of the child's reality.* Cambridge, England: Cambridge University Press.

Sullivan, P. M., & Burley, S. (1999). Mental testing of the hearing impaired child. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Intelligence and achievement* (pp.761-788). New York: Guilford.

Syed, N. (1993). *Development of a Non Verbal Test of Intelligence for Pakistani Urban Primary School Children.* Unpublished M. Phil. thesis, National Institute of Psychology, Quaid-i-Azam University, Islamabad, Pakistan.

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

Thompson, B., & Vacha-Haase, T. (2000). Psychometrics is datametrics: The test is not reliable. *Educational and Psychological Measurement, 60*, 174-195.

Thompson, L. A., Detterman, D. K., & Plomin, R. (1991). Association between cognitive abilities and scholastic achievement: Genetic overlap but environmental differences. *Psychological Science, 2*, 158-165.

Thorndike, E. L. (1924). The measurement of intelligence: Present status. *Psychological Review, 31*, 219-252.

Thorndike, R. L., & Hagen, E. P. (1987). *Cognitive Abilities Tests: Technical manual.* Chicago: Riverside Publishing Company.

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

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

Thorndike, R. L., & Hagen, E. P. (1991). *Measurement and evaluation in psychology and education* (5th ed.). New York: Macmillan.

Thurstone, L. L. (1938). Primary mental abilities. *Psychometric Monograph*, No. 1, Chicago: University of Chicago Press.

Ulosevich, S. N., Michael, W. B. & Bachelor, P. (1991). Higher-order factors in Structure of Intellect (SOI) aptitude tests hypothesised to portray construct of military leadership: A re-analysis of an SOI data base. *Educational and Psychological Measurement, 51*, 15-37.

Undheim, J. O. (1976). Ability structure in 10-11-year-old children and the theory of fluid and crystallized intelligence. *Journal of Educational Psychology, 68*, 411-423.

Undheim, J. O. (1978). Broad ability factors in 12-13- years-old children, the theory of fluid and crystallized intelligence, and the differentiation hypothesis. *Journal of Educational Psychology*, *70*, 433-443.

Undheim, J. O. (1981). On intelligence IV: Toward a restoration of general intelligence. *Scandinavian Journal of Psychology*, *22*, 251-265.

Undheim, J., & Gustafsson, J. (1987). The hierarchical organization of cognitive abilities: Restoring general intelligence through the use of linear structural relations. *Multivariate Behavioural Research*, *22*, 149-171.

United States Office of Personnel Management and Educational Testing Service. (1980). *Construct validity in psychological measurement: Proceedings of a Colloquium on Theory and Application in Education and Employment*. Princeton, NJ: Educational Testing Service.

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

Vernon, P. E. (1961). *The measurement of abilities*. London: University of London Press.

Vernon, P. E. (1965). Ability factors and environmental influence. *American Psychologist*, *9*, 723-733.

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

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

Vernon, P. A. (1987). *Speed of information processing and intelligence*. Norwood, NJ: Ablex.

Vernon, P. A. (1993). *Biological approach to the study of human intelligence.* Norwood, NJ: Ablex.

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

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

Wechsler, D. (1981). *Manual for Wechsler Adult Intelligence Scale-Revised.* New York: Harcourt.

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

Weinberg, R. A. (1989). Intelligence and IQ: Landmark issues and great debates. *American Psychologist, 44 (2)*, 94-104.

Wigdor, A. K. & Garner, W, R. (1982). *Ability testing: Uses, consequences, and controversies*, Part 11: Documentation section. Washington: National Academy Press.

Willard, L. S. (1968). A comparison of culture fair test scores with group and individual intelligence test scores of disadvantaged Negro children. *Journal of Learning Disabilities, 1*, 584-589.

Williams, R.l. (1974). Scientific racism and IQ: the silent mugging of the black community. *Psychology Today, 7* (12), 32-41.

Willis, S. L., & Nesselroade, C. S. (1990). Long-term effects of fluid ability training in old-old age. *Developmental Psychology, 26*, 905-910.

Womer, M. (1972). Culture and the concept of intelligence: A case in Uganda. *Journal of Cross-Cultural Psychology, 3,* 327-328.

ANNEXURES

ANNEXURE-A

Institutions/Centres - Sample for Second Tryout

1. Army Selection and Recruitment Centre Rawalpindi
2. Army Selection and Recruitment Centre Lahore
3. Army Selection and Recruitment Centre Peshawar
4. Army Selection and Recruitment Centre Karachi

ANNEXURE-B

Institutions/Centres - Sample for Final Study

1. Army Selection and Recruitment Centre Rawalpindi
2. Army Selection and Recruitment Centre Peshawar
3. Army Selection and Recruitment Office Sargodha
4. Army Selection and Recruitment Centre Lahore
5. Army Selection and Recruitment Centre Karachi
6. Army Selection and Recruitment Centre Quetta
7. Inter Services Selection Board Kohat
8. Army Public School and College for Boys, Rawalpindi
9. Army Public School and College for Girls, Rawalpindi
10. Fatima Jinnah Women University (Under Graduate Campus) Rawalpindi.

ANNEXURE-C

Institutions/Centres/Offices- Standardization Sample

1. Army Selection and Recruitment Centre Rawalpindi
2. Army Selection and Recruitment Centre Lahore
3. Army Selection and Recruitment Office Sialkot
4. Army Selection and Recruitment Office Faisal Abad
5. Army Selection and Recruitment Centre Multan
6. Army Selection and Recruitment Office Sargodha
7. Army Selection and Recruitment Centre Karachi
8. Army Selection and Recruitment Centre HyderAbad
9. Army Selection and Recruitment Office Sukkar
10. Army Selection and Recruitment Centre Quetta
11. Army Selection and Recruitment Office Kohat
12. Army Selection and Recruitment Centre Peshawer
13. Army Public School and College for Boys, Rawapindi
14. Inter Services Selection Board Kohat
15. Inter Services Selection Board Gujranwala

Annexure-D

ANSWER SHEET

NAME : ______________________ NO : ______________________

AGE : ______________________ CLASS : ______________________

SCHOL/COLLEGE : ______________________ RURAL/URBAN: ______________________

MEDIUM OF INSTRUCTION ______________________ MARKS : ______________________

(MATRIC) (INTERMEDIATE/1st YEAR)

SCIENCE/ARTS ______________________

SERIES	MATRICES	ANALOGIES	ODD-ONE OUT	SIMILARITIES

SERIES

S/NO	A
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	

MATRICES

S/NO	A
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	

ANALOGIES

S/NO	A
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	

ODD-ONE OUT

S/NO	A
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	

SIMILARITIES

S/NO	A
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	

GENERAL INSTRUCTIONS- ہدایات

1. This is a nonverbal Intelligence Test in which different figures and patterns have been used instead of words.

۱۔یہ ایک ذہانت کا ٹیسٹ ہے۔جس میں الفاظ کی بجائے مختلف اشکال اور خاکے وغیرہ استعمال کیئے گئے ہیں۔

2. For this test, you will be provided a booklet and a answer sheet. First of all, you will complete the particulars given on the answer sheet. Do not open teh booklet until you are total to do so.

۲۔اس ٹیسٹ کیلئے آپکو ایک کتابچہ اور ایک جوابی پرچہ دیا جائے گا۔سب سے پہلے آپ کو جوابی پرچہ پر دیئے گئے کوائف پُر کرنے ہیں۔کتابچے کو اس وقت تک نہ کھولیں، جب تک آپ کو کہا نہ جائے۔

3. This test consist of five subtests. At the start of each test, some instructions are given. Please reead these instructions carefully.

۳۔یہ ٹیسٹ پانچ حصوں پر مشتمل ہے۔ہر ٹیسٹ کے شروع میں اس کو حل کرنے کیلئے ہدایات دی گئی ہیں۔ٹیسٹ شروع کرنے سے پہلے ان ہدایات کو غور سے پڑھیں۔

4. All the tests have been explained with the help of examples. Do these examples carefully.

۴۔پانچوں ٹیسٹوں کے شروع میں ہدایات کے بعد کچھ مثالیں اور ان کے جواب بھی دیئے گئے ہیں۔آپ ان کا بغور جائزہ لیں۔

5. According to peculiar characteristics of these tests, number of answer options of each test range from 4 to 6. There is only one correct answer to each question. All the answers are numbered. Choose teh correct answer and write its number on teh answer sheet against the question number.

۵۔ان ٹیسٹوں کی نوعیت کے لحاظ سے کسی ٹیسٹ میں ہر سوال کے چار کسی میں پانچ اور کسی میں چھ ممکنہ جوابات دیئے گئے ہیں۔جن میں صرف ایک جواب درست ہے۔تمام ممکنہ جوابات کو ایک نمبر دیا گیا ہے آپ نے اس درست جواب کا انتخاب کر کے اسے جوابی پرچہ میں اس سوال کے سامنے والے خانے میں لکھنا ہے۔

6. If you want to change any answer, draw a cross (x) on it and write the correct answer in the same column.

٦۔اگر آپ کسی جواب کو تبدیل کرنا چاہیں تو اس جواب پر کراس (X) کا نشان لگا دیں اور دوسرا جواب اس خانے میں لکھ دیں۔

Annexure-F

SERIES

Time limit – 10 Minutes

INSTRUCTIONS

Each item of this test consists of two sets of figures. Figures on the left side go together to form some kind of series and are called probelm figures. Figures on the right side are called answer figures. Each one of the answer figures is numbered. Find out which one of the answer figure that goes where you see the question mark (?) in the series. Write its number on the answer sheet.

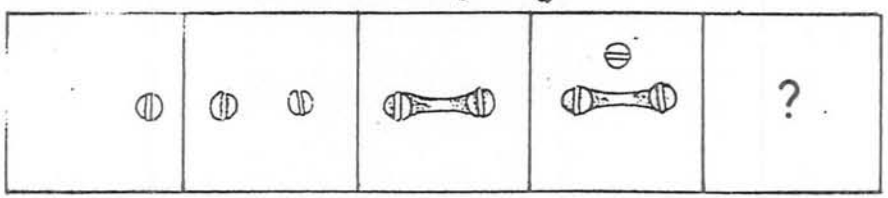
ہدایات

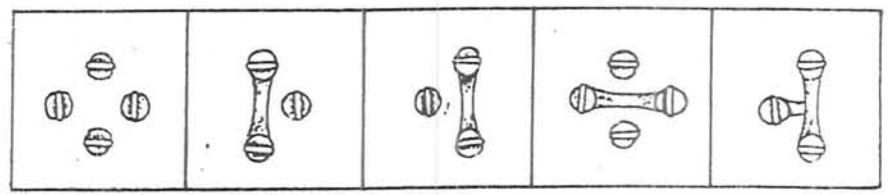
اس ٹیسٹ کے سوالات دو حصوں پر مشتمل ہیں۔ ایک حصے میں سوالیہ اشکال اور دوسرے حصے
میں جوابی اشکال ہیں۔ آپ نے ہر سوال کی جوابی اشکال میں سے اس شکل کا انتخاب کرنا ہے
جس کو سوالیہ اشکال میں دیئے گئے سوالیہ نشان (؟) کے خانے میں آنا چاہیے اس شکل کے نیچے
جو نمبر دیا گیا ہے اس نمبر کو جوابی فارم میں اس سوال کے سامنے والے خانے میں لکھ دیں۔

EXAMPLES

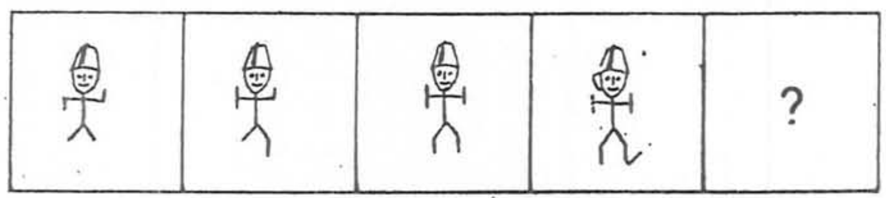
PROBLEM FIGURES سوالیہ اشکال

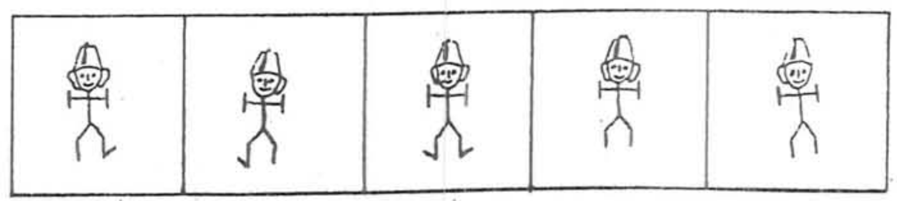
ANSWER FIGURES جوابی اشکال

1. 



1 2 3 4 5

2. 

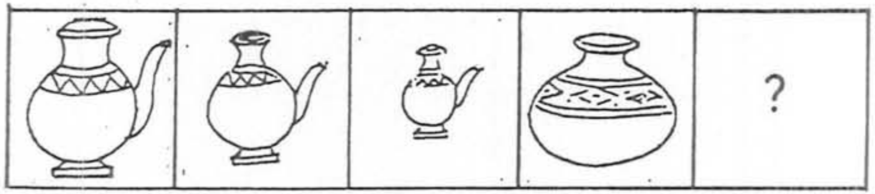


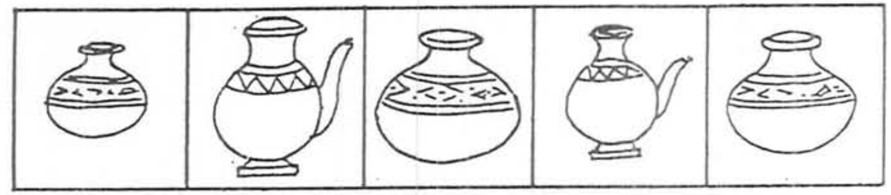
1 2 3 4 5

PROBLEM FIGURES سوالیہ اشکال

ANSWER FIGURES جوابی اشکال

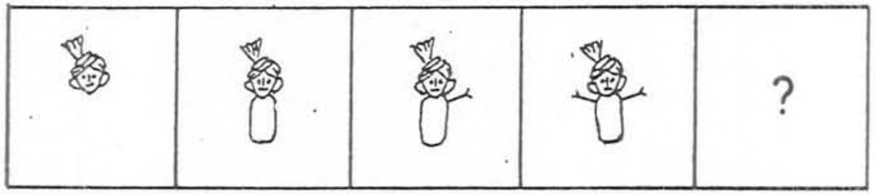
1.

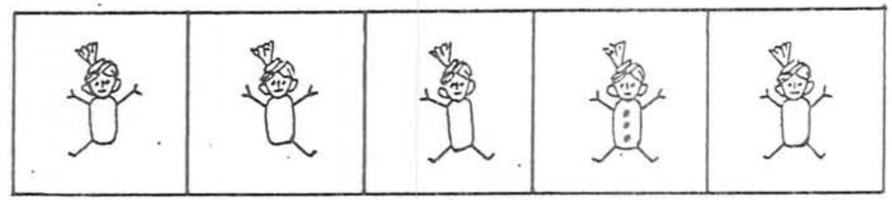




1 2 3 4 5

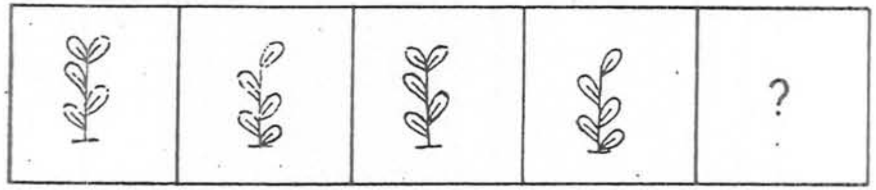
2.

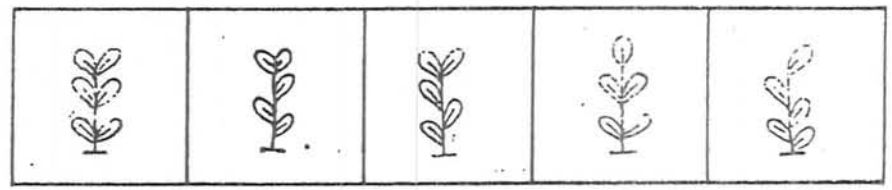




1 2 3 4 5

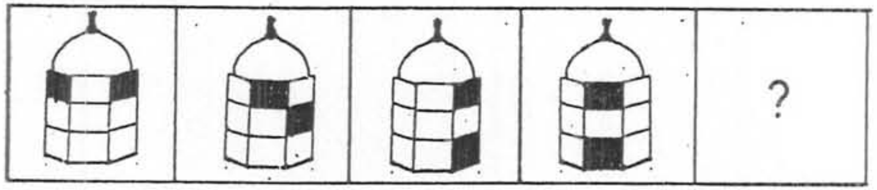
3.

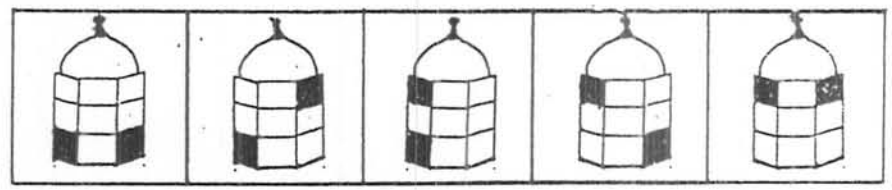




1 2 3 4 5

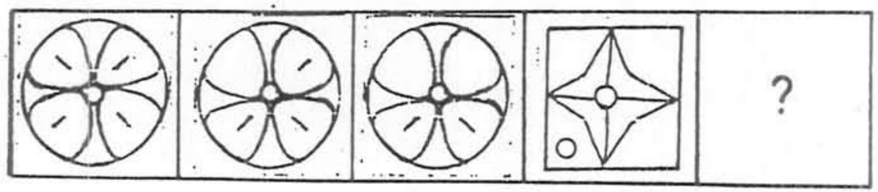
4.

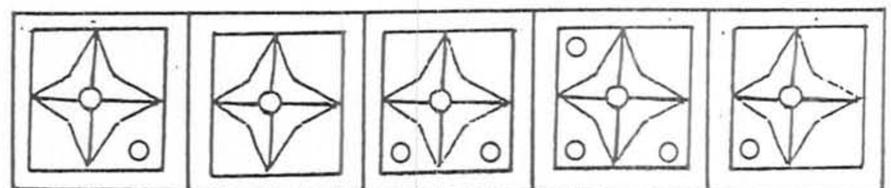




1 2 3 4 5

5.



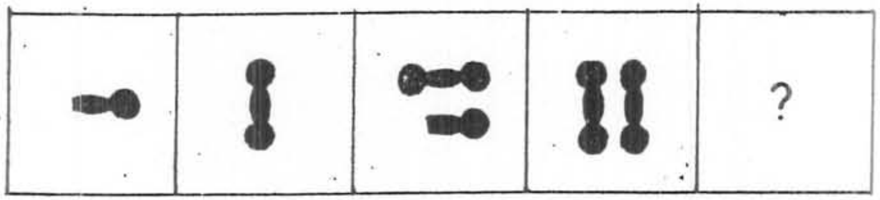


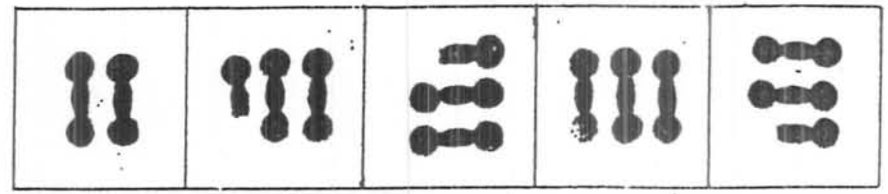
1 2 3 4 5

PROBLEM FIGURES سوالیہ اشکال

ANSWER FIGURES جوابی اشکال

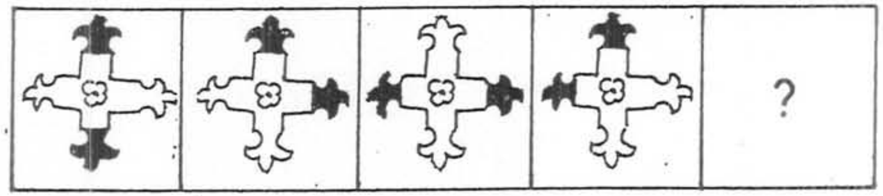
6.

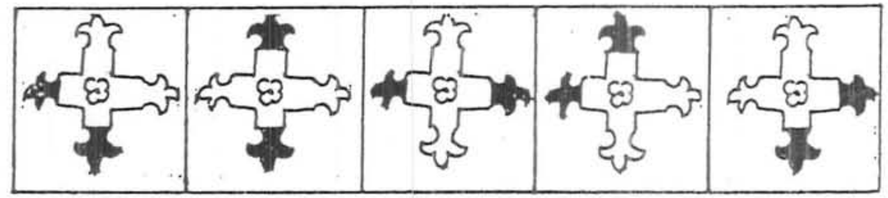




1 2 3 4 5

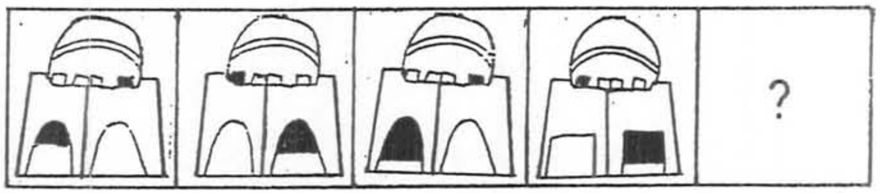
7.

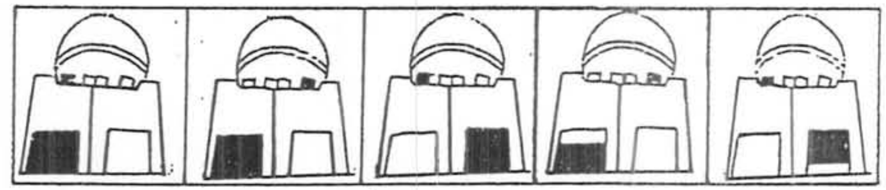




1 2 3 4 5

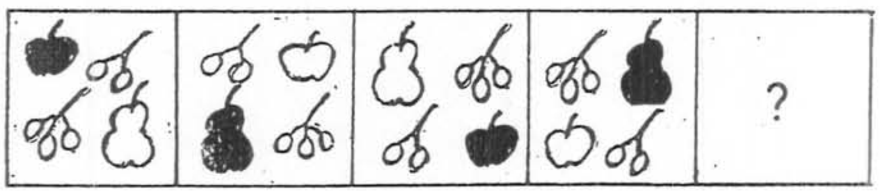
8.

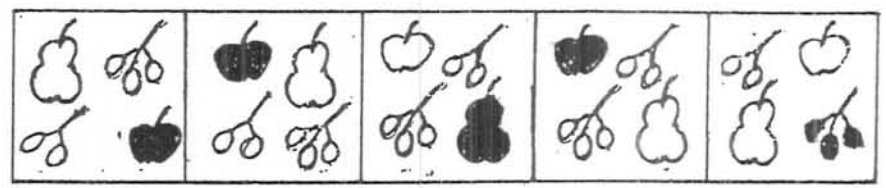




1 2 3 4 5

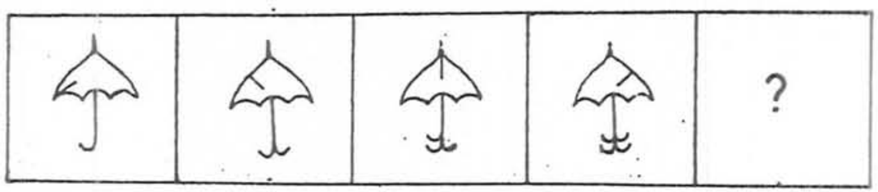
9.

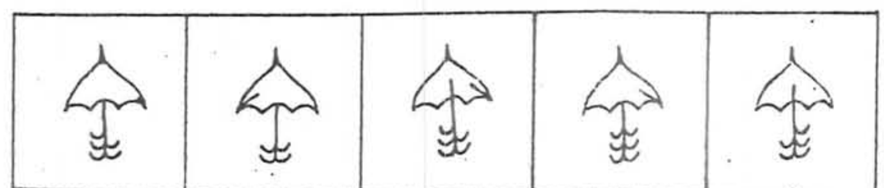




1 2 3 4 5

10.

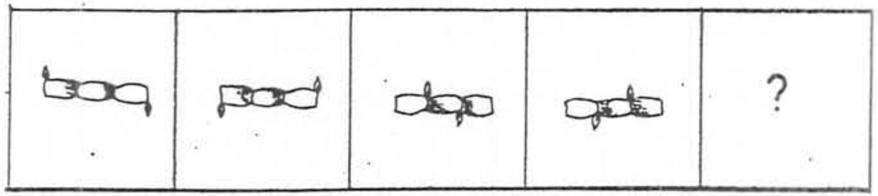


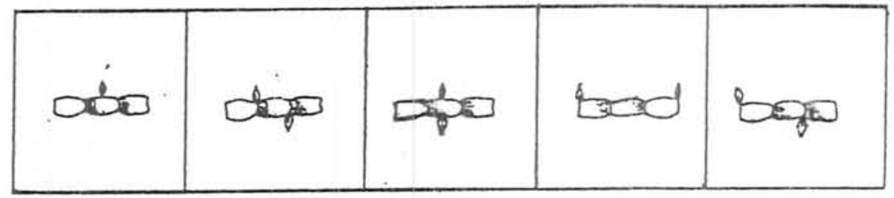


1 2 3 4 5

PROBLEM FIGURES سوالیہ اشکال — ANSWER FIGURES جوابی اشکال

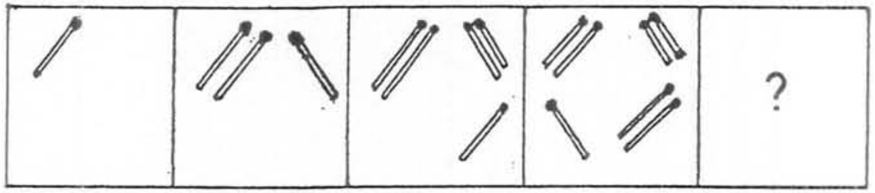
11.

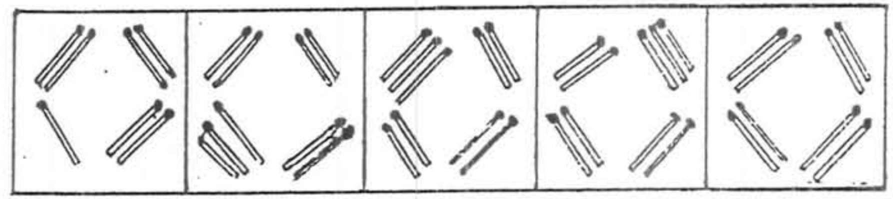




1 2 3 4 5

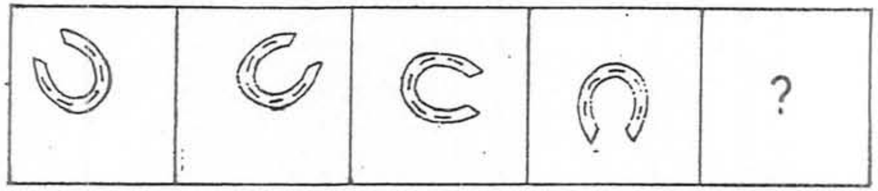
12.

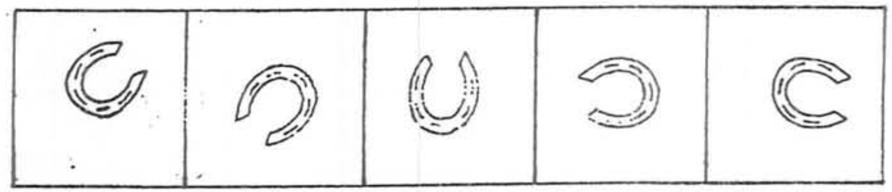




1 2 3 4 5

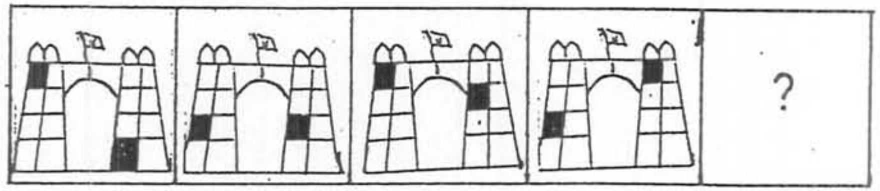
13.

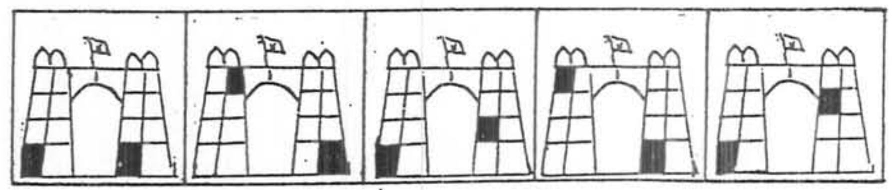




1 2 3 4 5

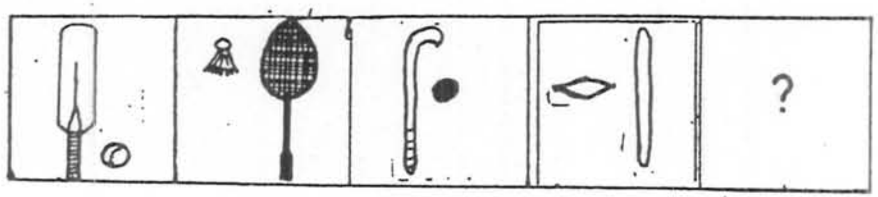
14.

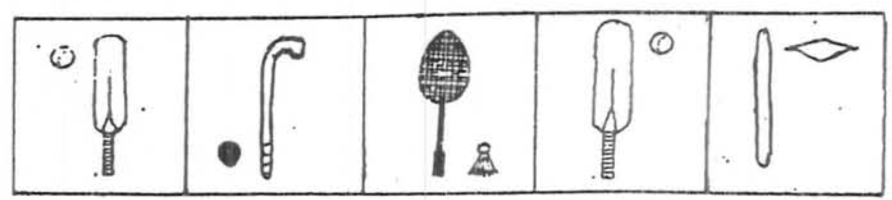




1 2 3 4 5

15



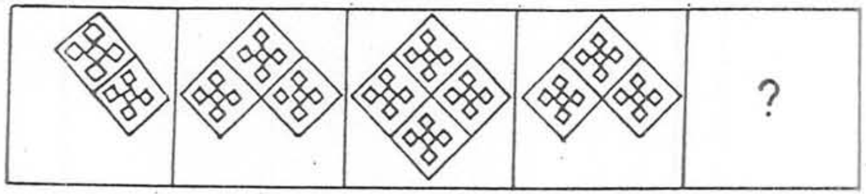


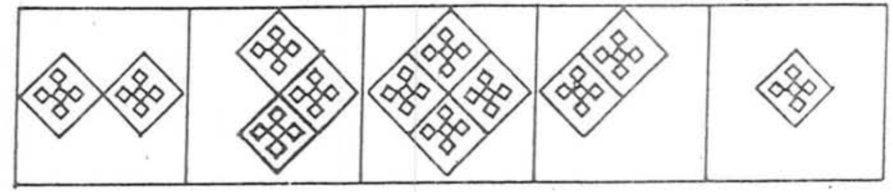
1 2 3 4 5

PROBLEM FIGURES سوالیہ اشکال

ANSWER FIGURES جوابی اشکال

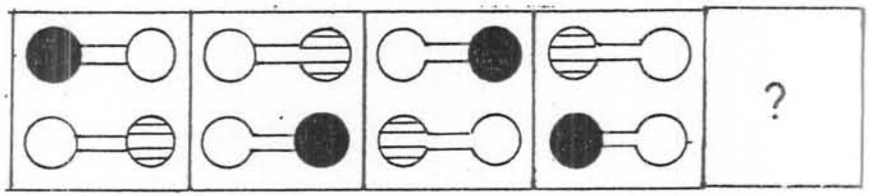
16.

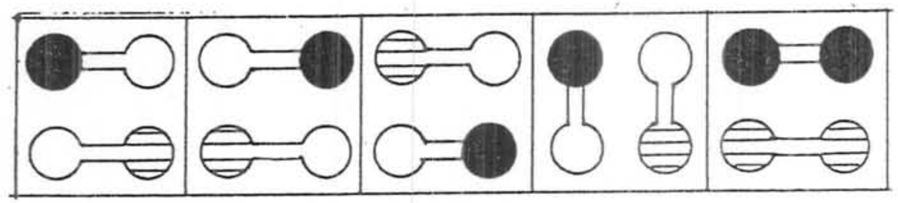




1 2 3 4 5

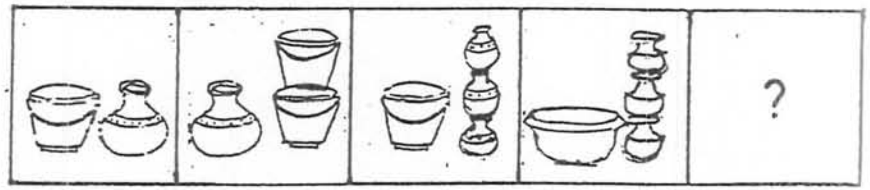
17.

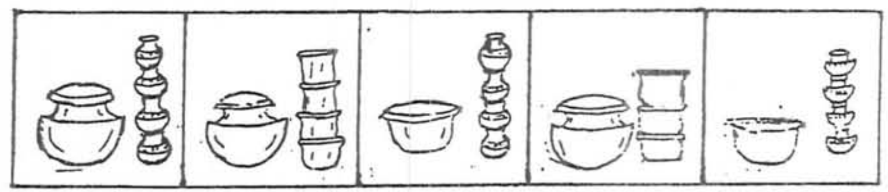




1 2 3 4 5

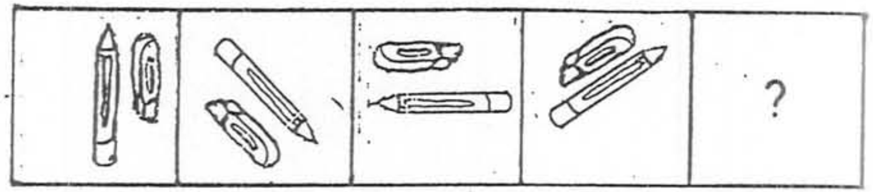
18.

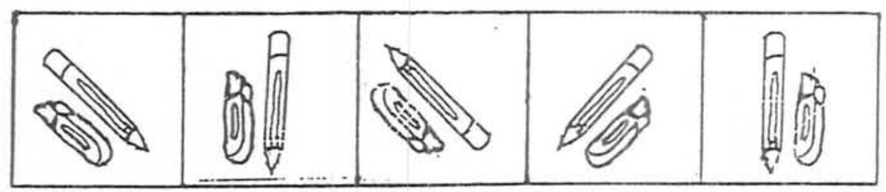




1 2 3 4 5

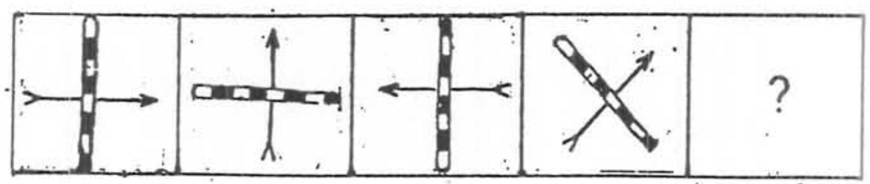
19.

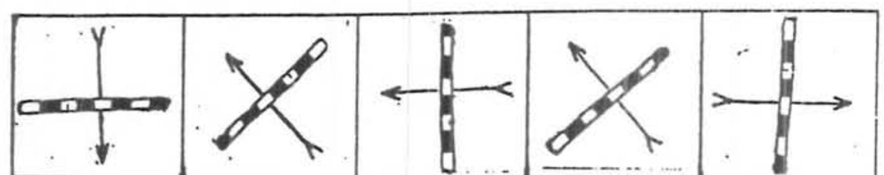




1 2 3 4 5

20.





1 2 3 4 5

MATRICES

Time limit – 10 Minutes

MATRICES

INSTRUCTIONS

Each item of this test consists of two sets of figures. Figures on the left side are called probelm figures and those on the right side are called answer figures. Each one of the answer figure is numbered. Find out which one of the answer figure that goes where you see the qestion mark (?). Write its number on the answer sheet.

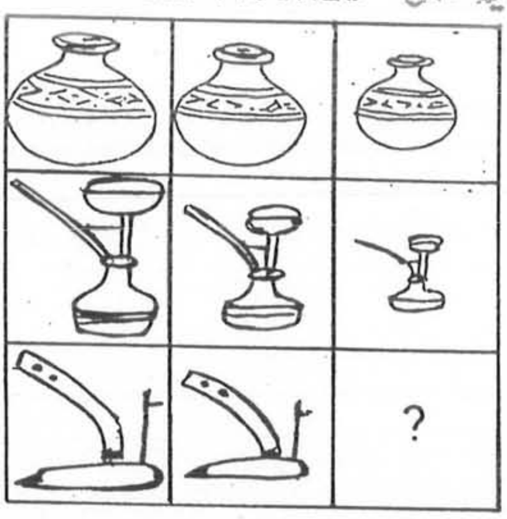
ہدایات

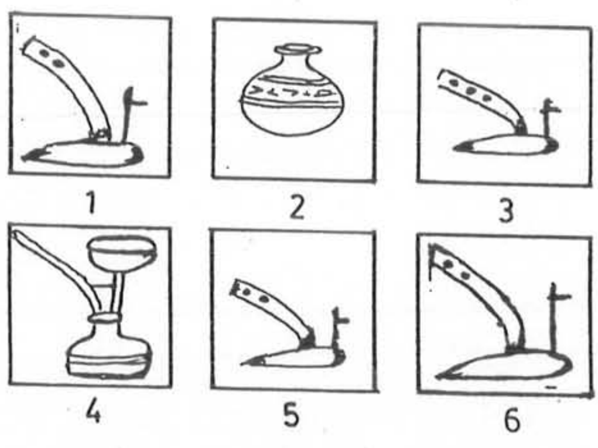
اس ٹیسٹ میں بائیں طرف والے حصے میں سوالیہ اشکال اور دائیں طرف جوابی اشکال ہیں۔ ہر سوال کے لئے آپ نے جوابی اشکال میں سے اس شکل کو تلاش کرنا ہے جس کو سوالیہ اشکال میں دیئے گئے سوالیہ نشان (؟) کی جگہ آنا چاہیے اس شکل کے نیچے جو نمبر دیا گیا ہے اس کو جوابی فارم میں اس سوال کے سامنے والے خانے میں لکھ دیں۔

EXAMPLES

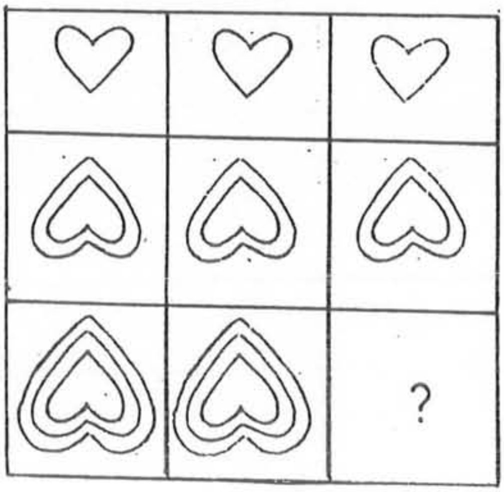
PROBLEM FIGURES سوالیہ اشکال — ANSWER FIGURES جوابی اشکال

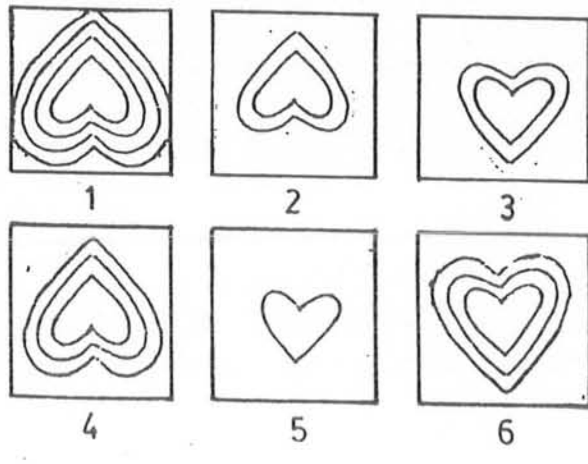
1.





2.

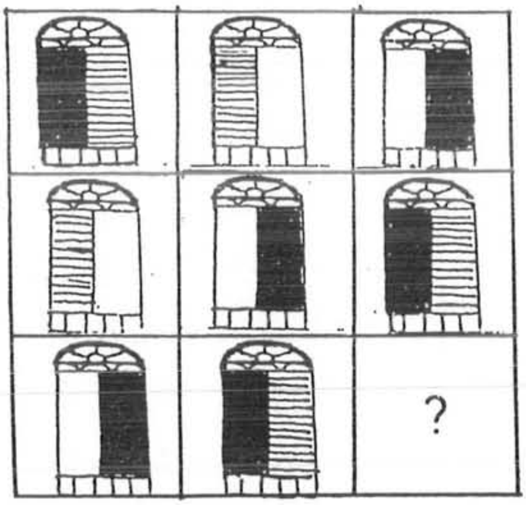


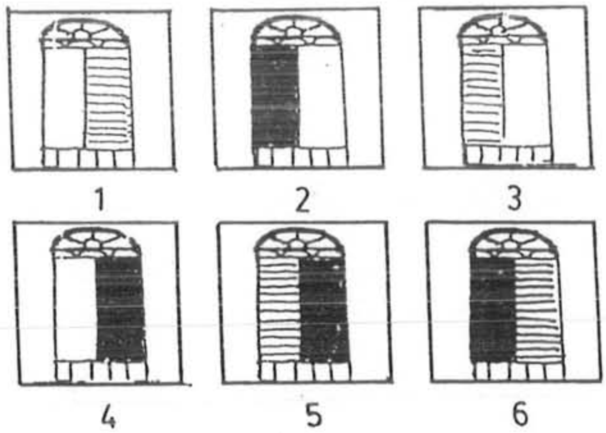


PROBLEM FIGURES سوالیہ اشکال

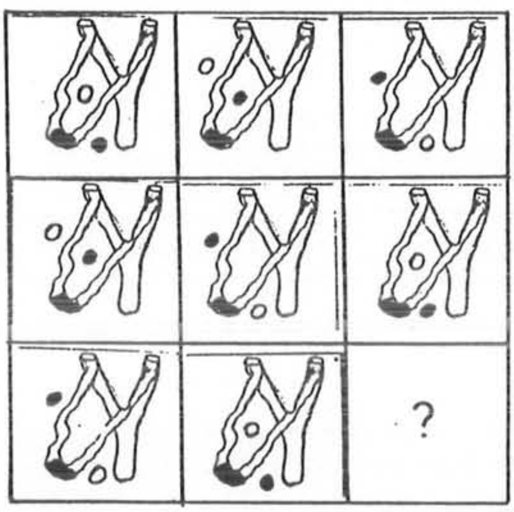
ANSWER FIGURES جوابی اشکال

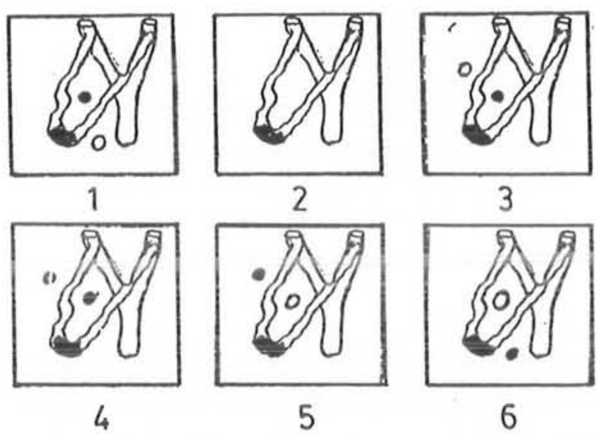
1.



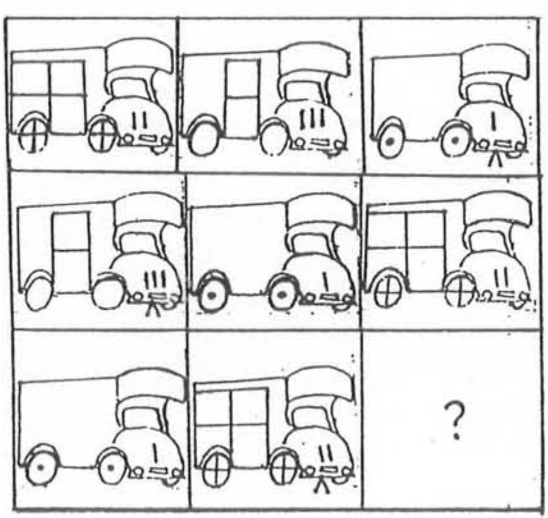


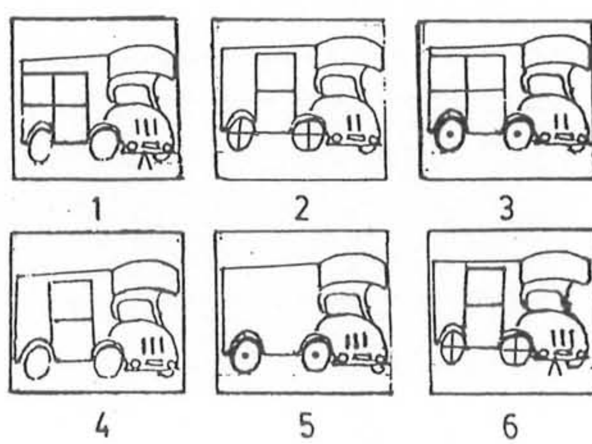
2.





3.

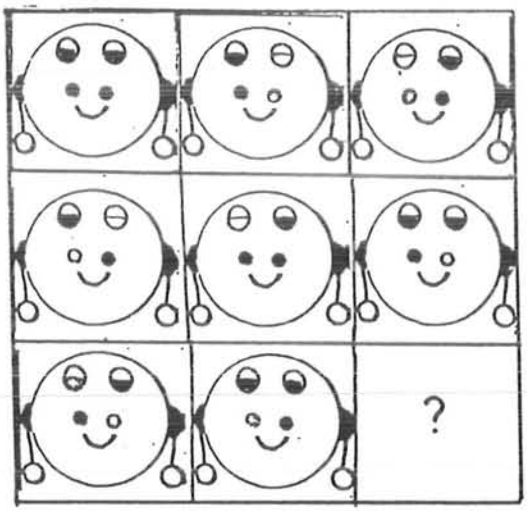


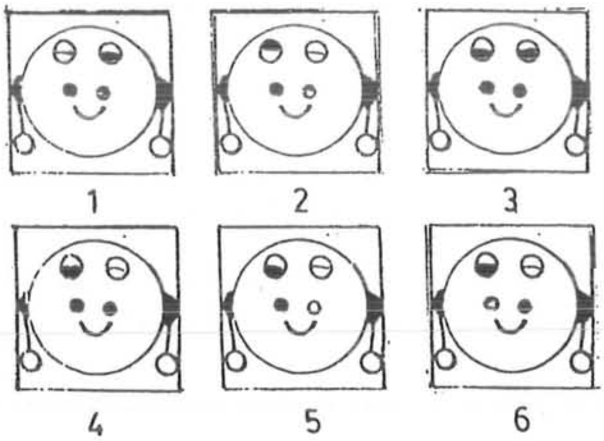


PROBLEM FIGURES سوالیہ اشکال

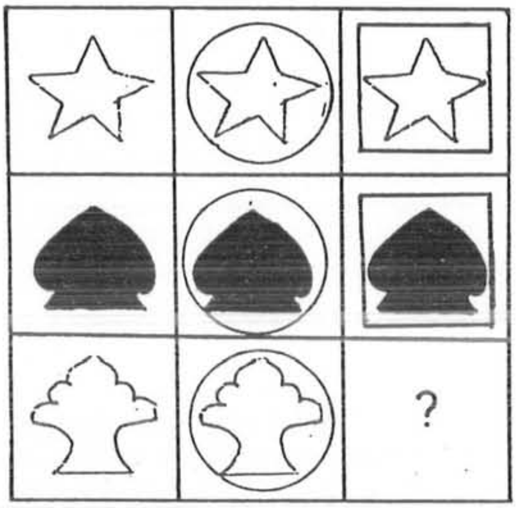
ANSWER FIGURES جوابی اشکال

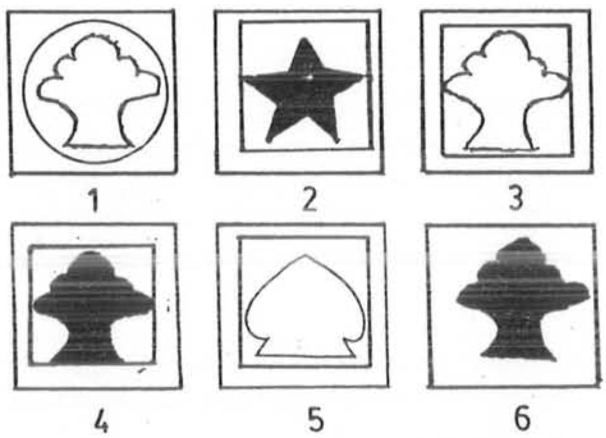
4.



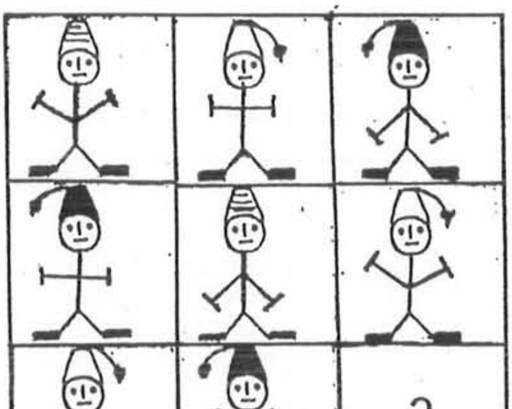


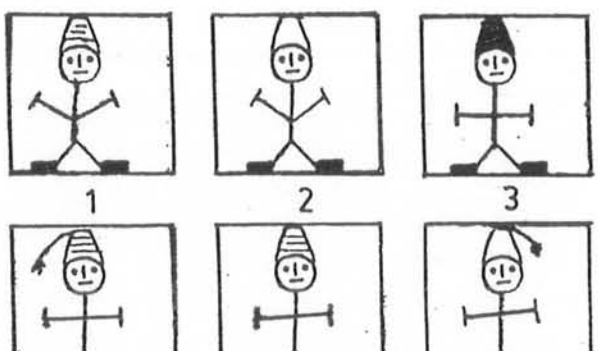
5.





6.

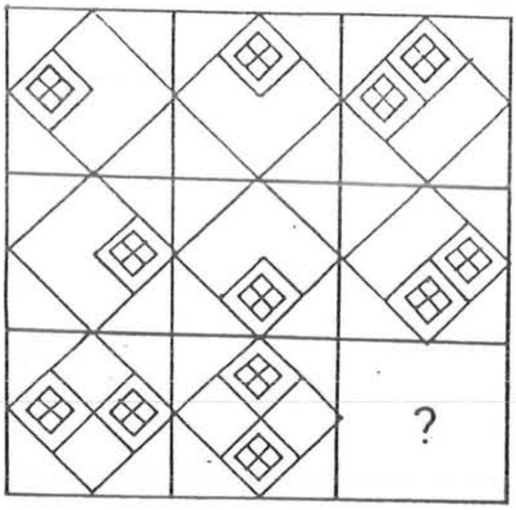


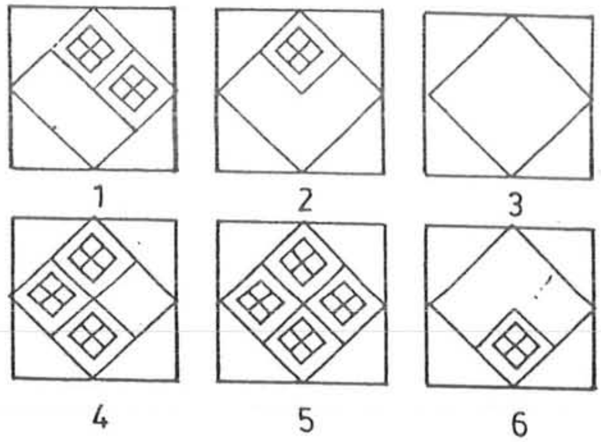


PROBLEM FIGURES سوالیہ اشکال

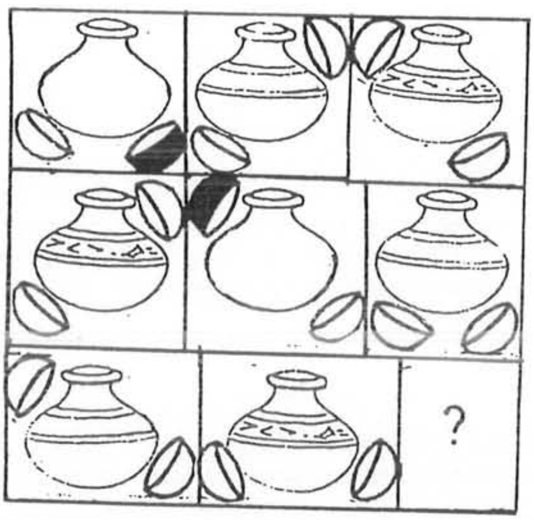
ANSWER FIGURES جوابی اشکال

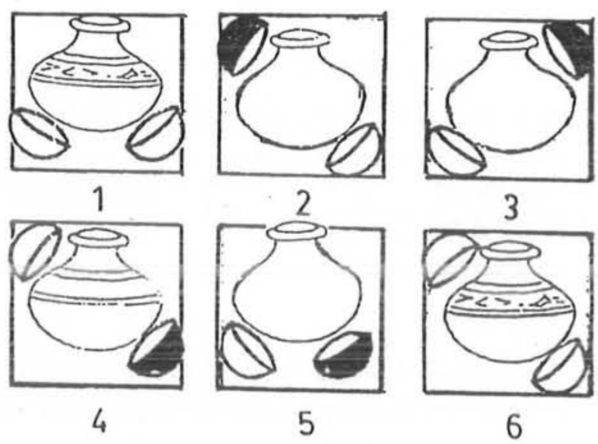
7.



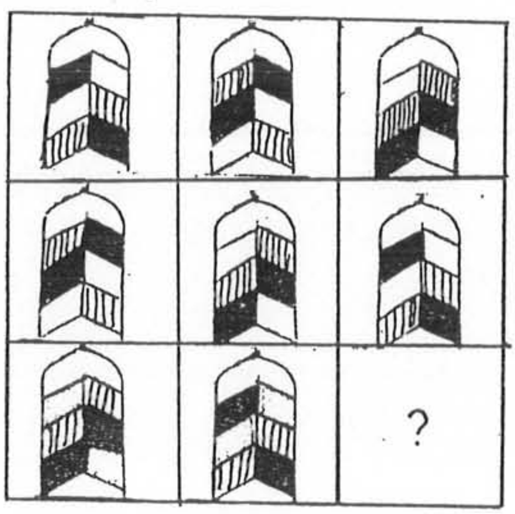


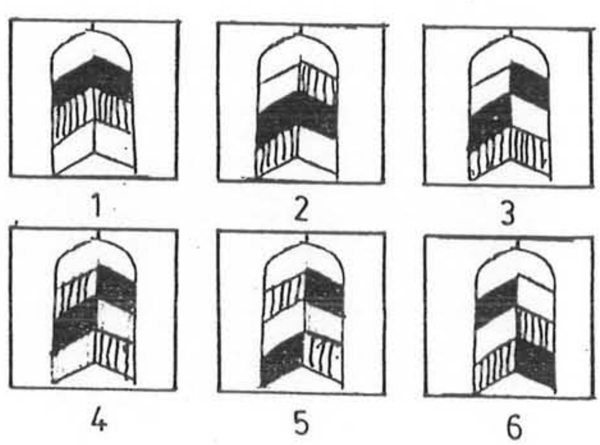
8.





9.

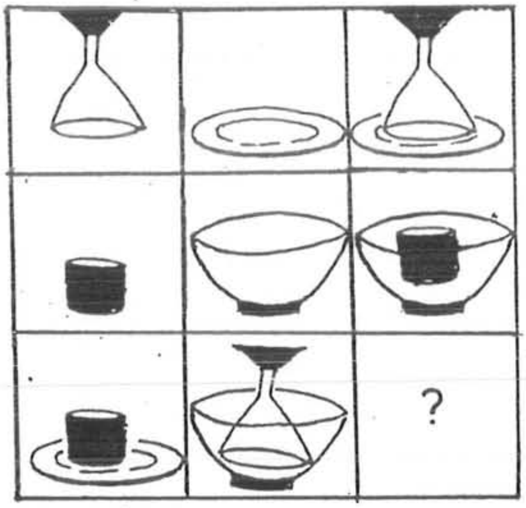


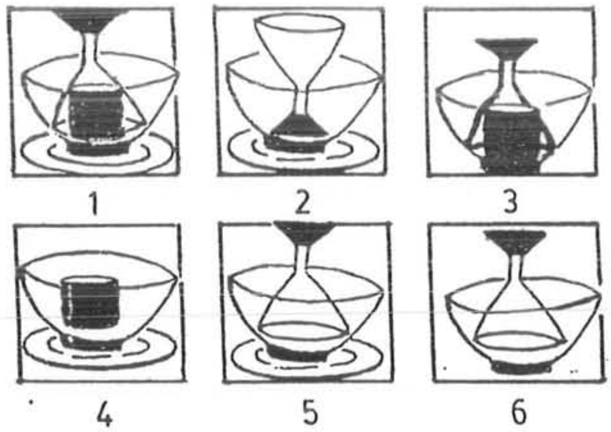


PROBLEM FIGURES سوالیہ اشکال

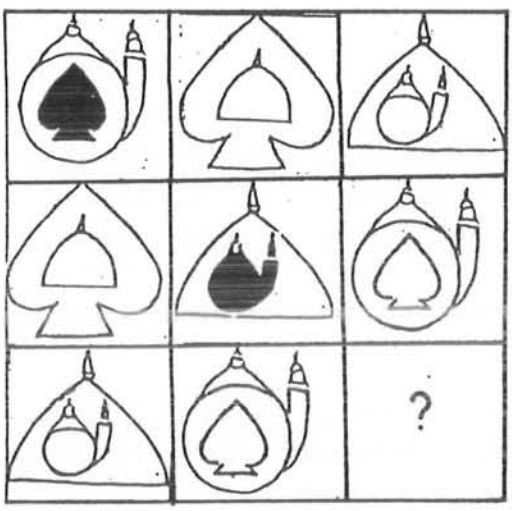
ANSWER FIGURES جوابی اشکال

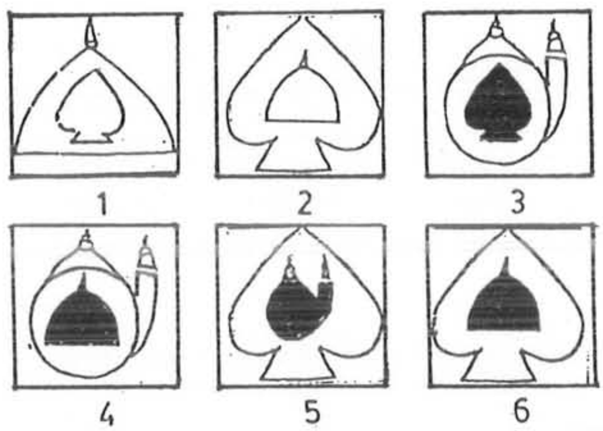
10.



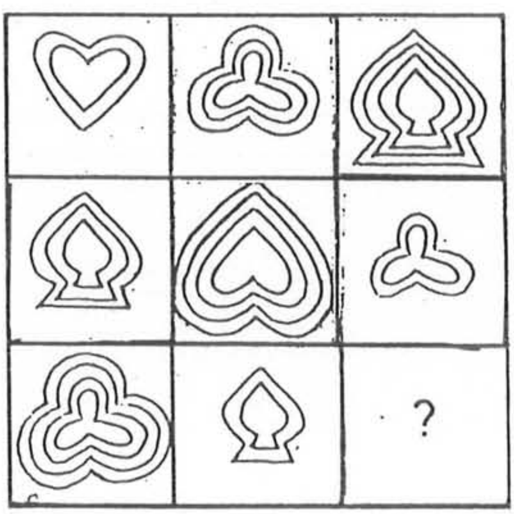


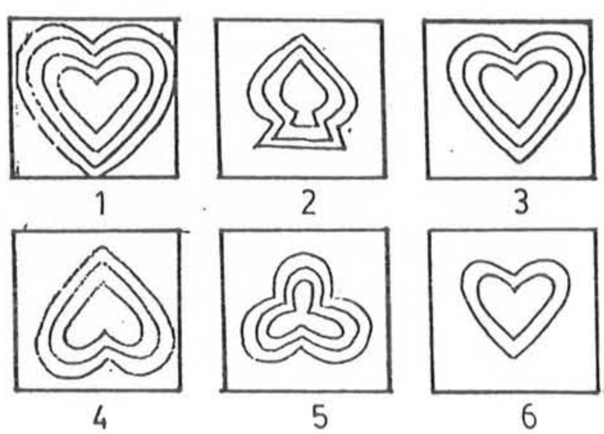
11.





12.

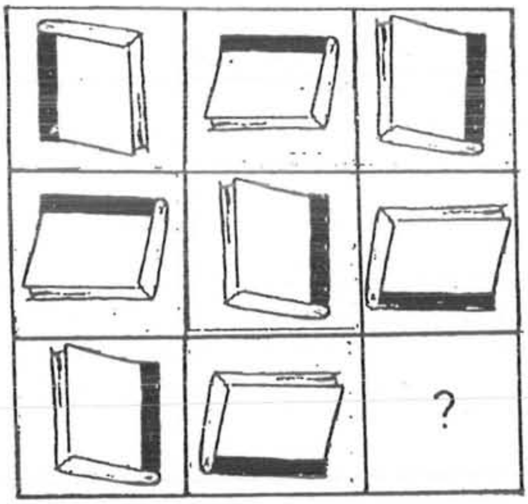


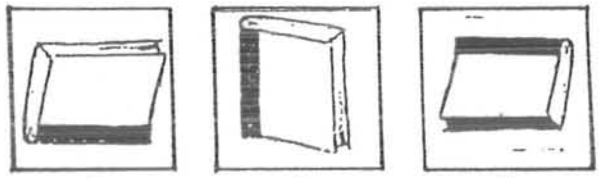


PROBLEM FIGURES سوالیہ اشکال

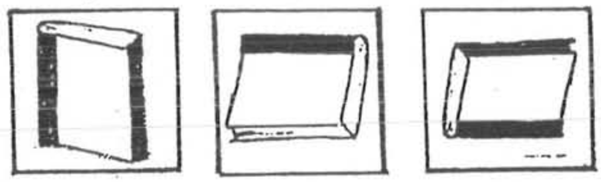
ANSWER FIGURES جوابی اشکال

13.



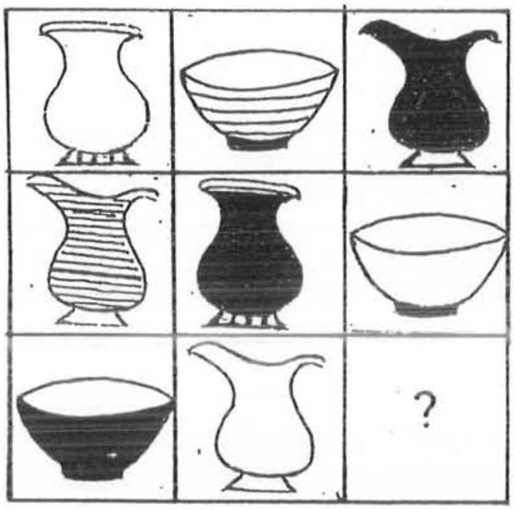


1 2 3



4 5 6

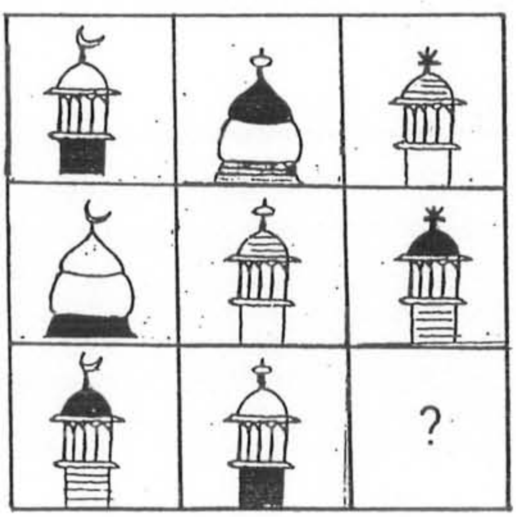
14.



1 2 3

4 5 6

15.



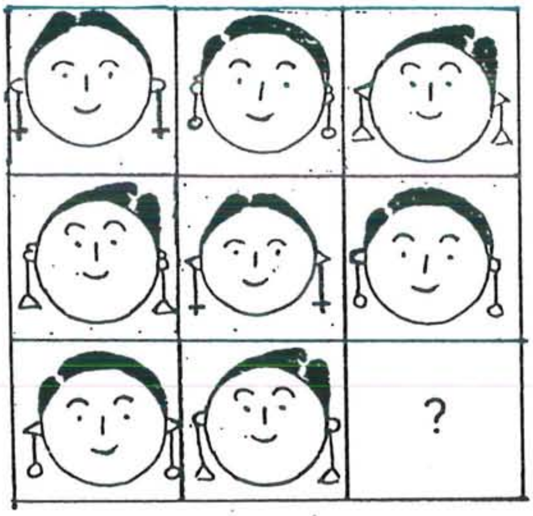
1 2 3

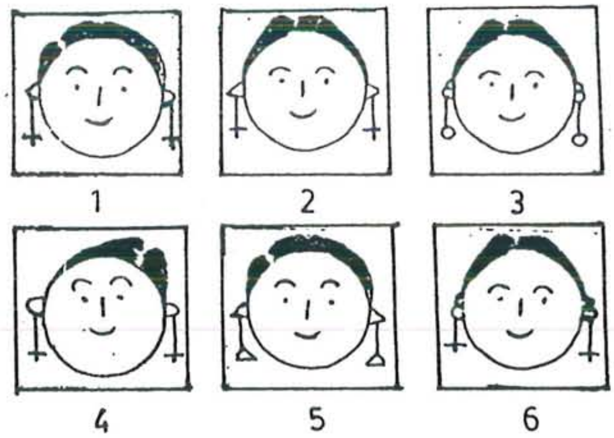
4 5 6

PROBLEM FIGURES سوالیہ اشکال

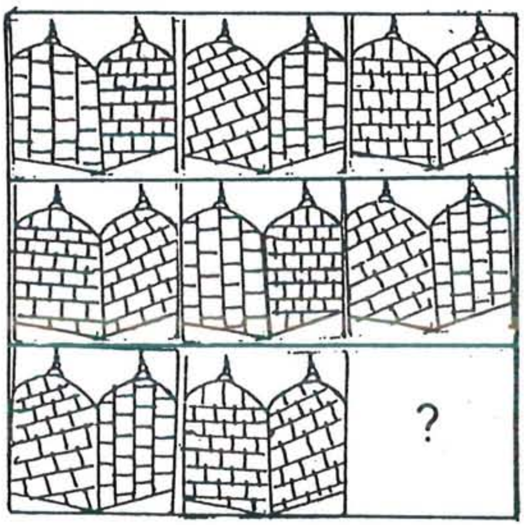
ANSWER FIGURES جوابی اشکال

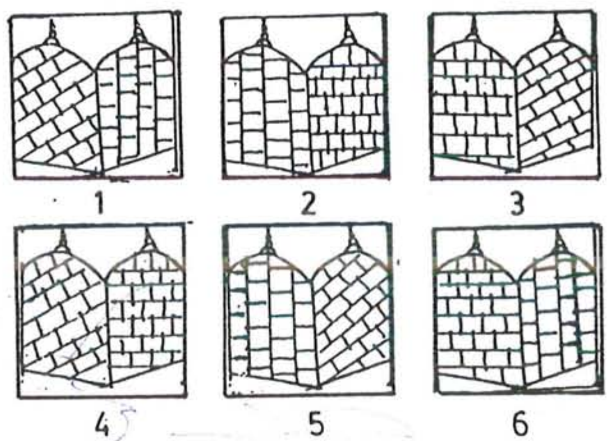
16.



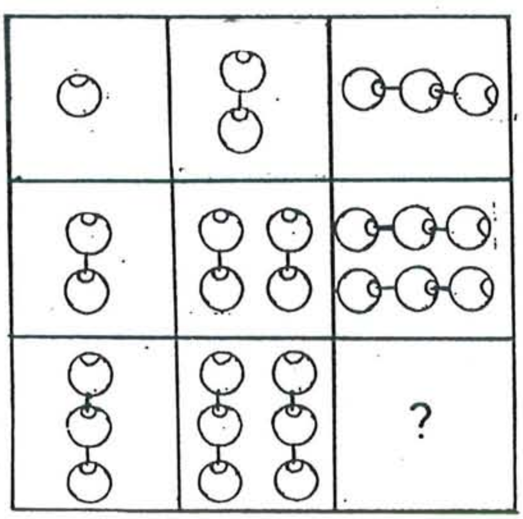


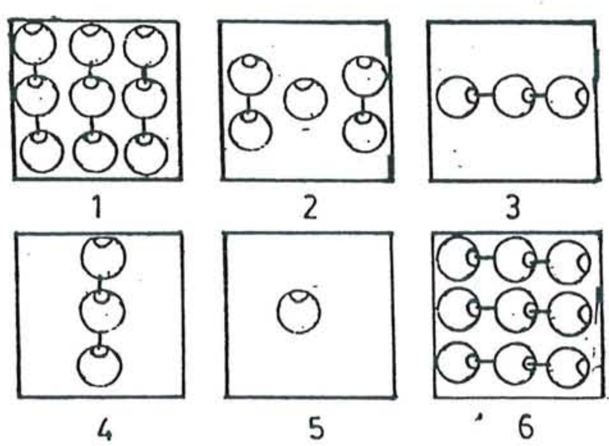
17





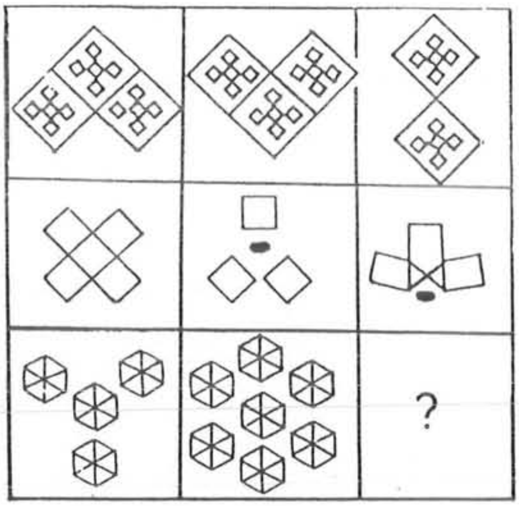
18.

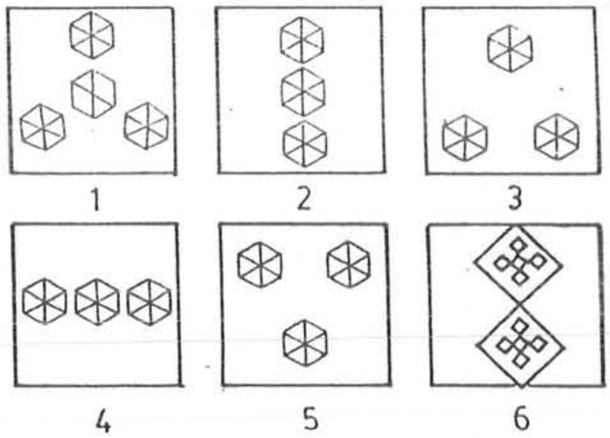


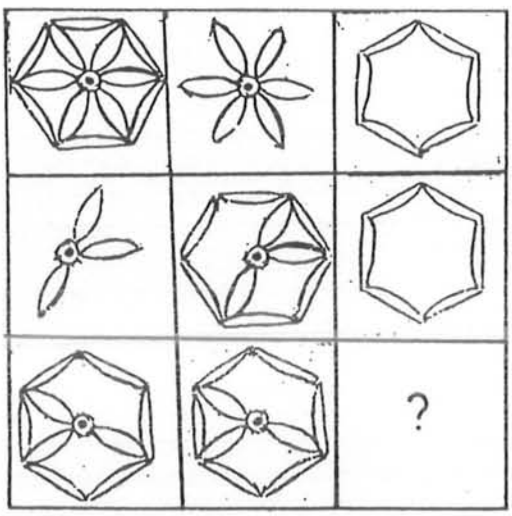


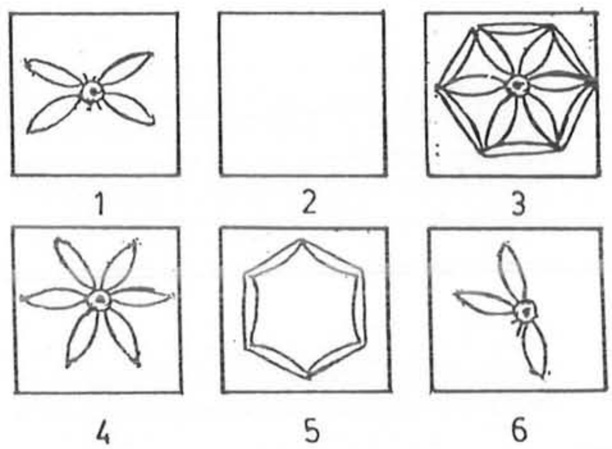
PROBLEM FIGURES سوالیہ اشکال

ANSWER FIGURES جوابی اشکال

19. 



20. 



ANALOGIES

Time limit – 10 Minutes

ANALOGIES

INSTRUCTIONS

Each item of this test consists of two sets of figures. Figures on the left side are called probelm figures and those on the right side are called answer figures. In problem figures there is a definite relationship between first and second figures. Find out the figure from the answer figures that goes with the third figure in the same way and write its number on the answer sheet.

ہدایات

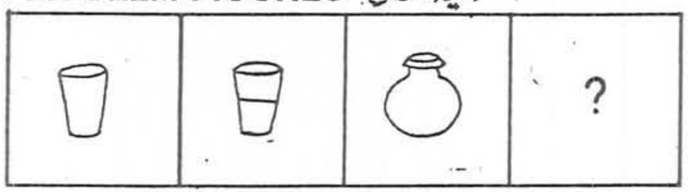
اس ٹیسٹ کے سوالات دو حصوں پر مشتمل ہیں۔ ایک حصے میں سوالیہ اشکال اور دوسرے حصے میں جوابی اشکال ہیں۔ سوالیہ اشکال میں پہلی اور دوسری شکل کے درمیان ایک واضح تعلق ہے اسی طرح کا تعلق تیسری شکل اور جوابی اشکال میں کسی ایک شکل کے درمیان ہے۔ آپ نے ہر سوال کی جوابی اشکال میں سے اس شکل کا انتخاب کرنا ہے جس کا تعلق سوالیہ اشکال میں دی گئی تیسری شکل سے ہو اور جس کو سوالیہ نشان (؟) کے خانے میں آنا چاہیے اس شکل کے نیچے جو نمبر دیا گیا ہے اس نمبر کو جوابی فارم میں اس سوال کے سامنے والے خانے میں لکھ دیں۔

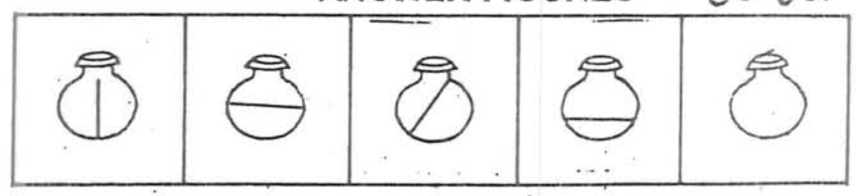
EXAMPLES

PROBLEM FIGURES سوالیہ اشکال

ANSWER FIGURES جوابی اشکال

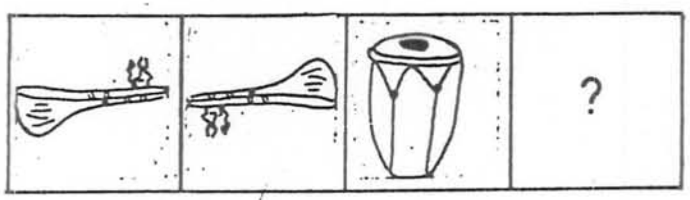
1.

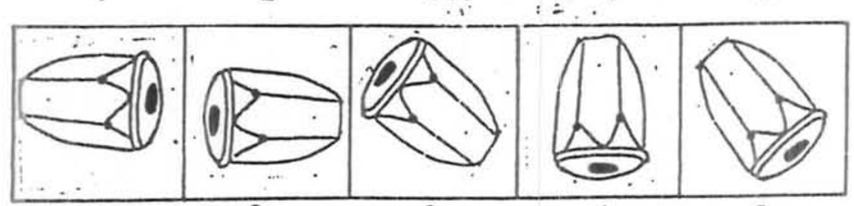




1 2 3 4 5

2.

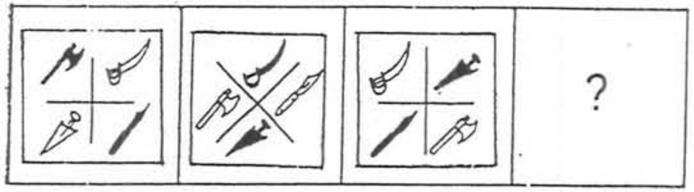


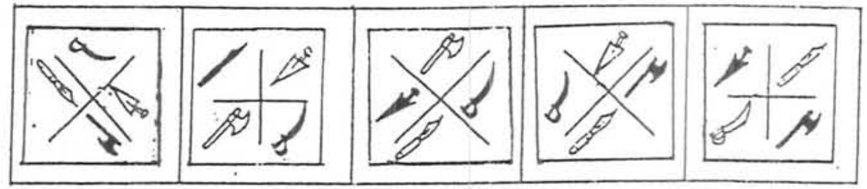


1 2 3 4 5

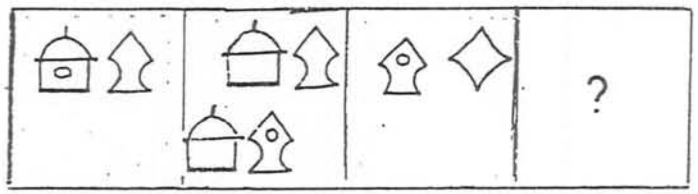
PROBLEM FIGURES سوالیہ اشکال

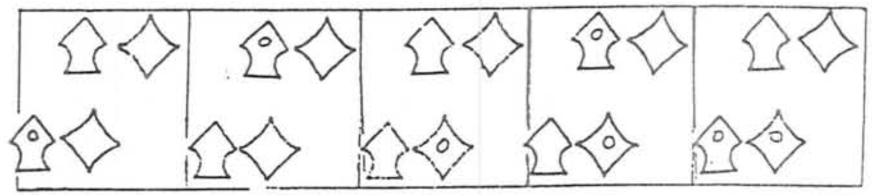
ANSWER FIGURES جوابی اشکال

1. 

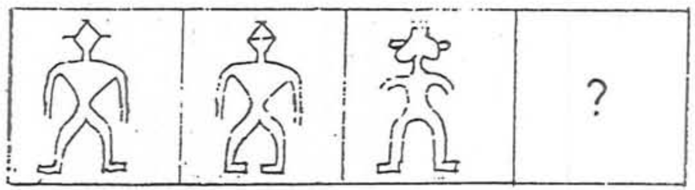


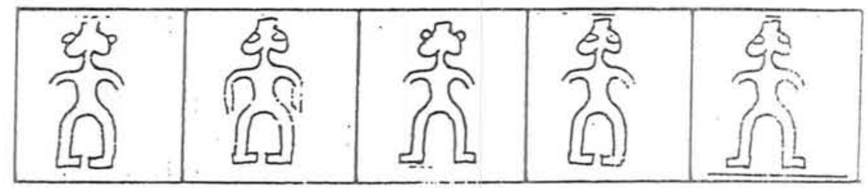
1 2 3 4 5

2. 

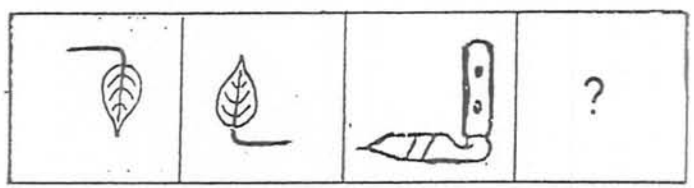


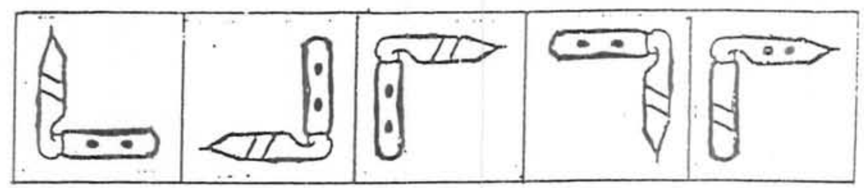
1 2 3 4 5

3. 

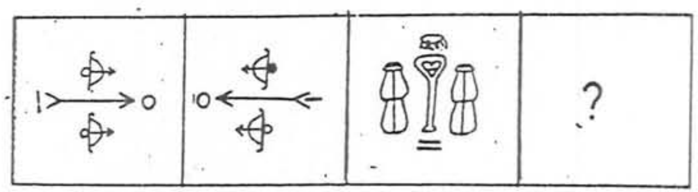


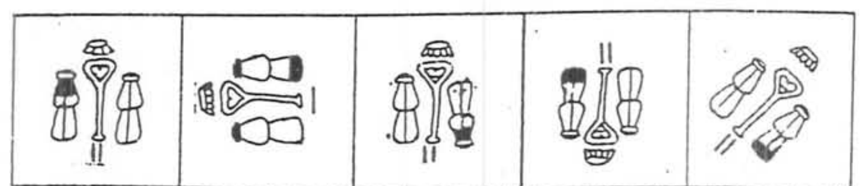
1 2 3 4 5

4. 



1 2 3 4 5

5. 

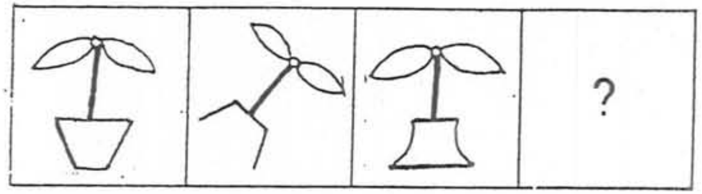


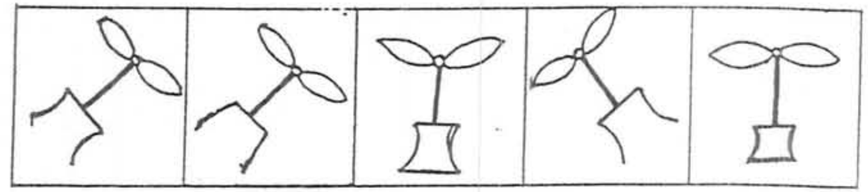
1 2 3 4 5

PROBLEM FIGURES سوالیہ اشکال

ANSWER FIGURES جوابی اشکال

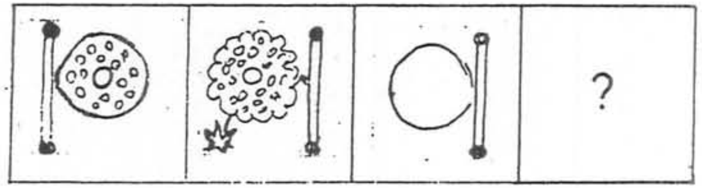
6.

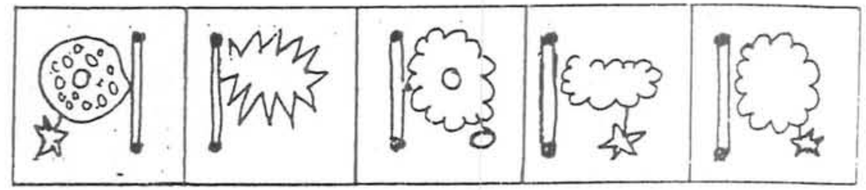




1 2 3 4 5

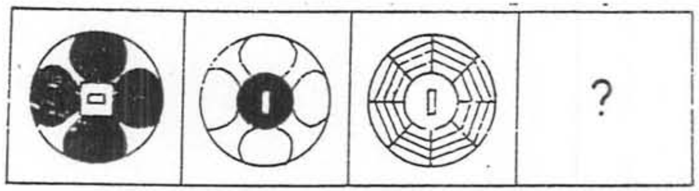
7.

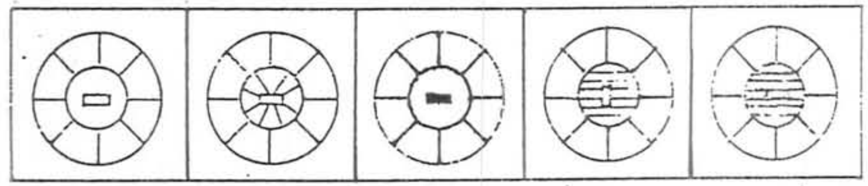




1 2 3 4 5

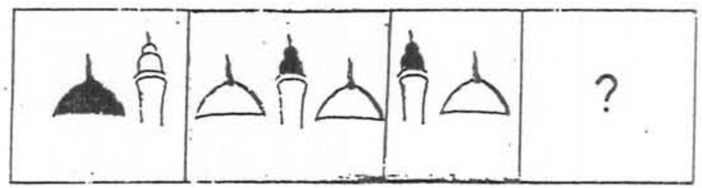
8.

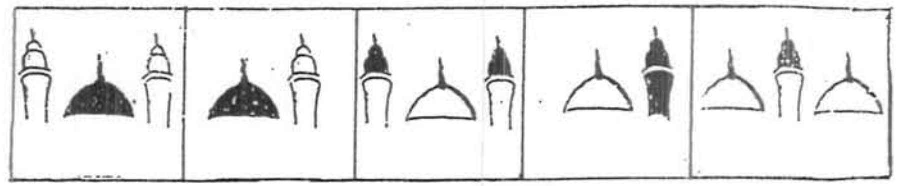




1 2 3 4 5

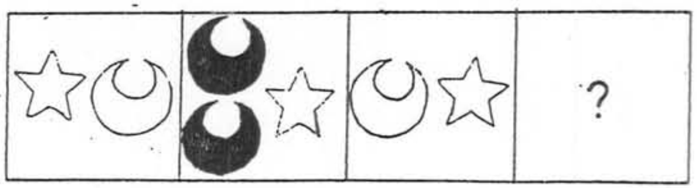
9.

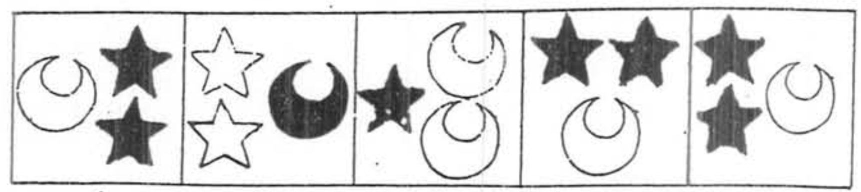




1 2 3 4 5

10.

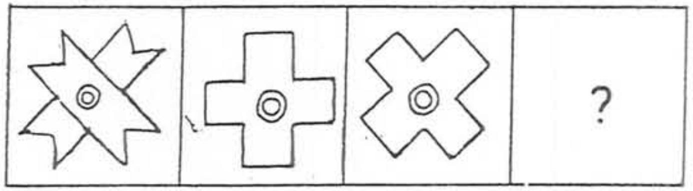
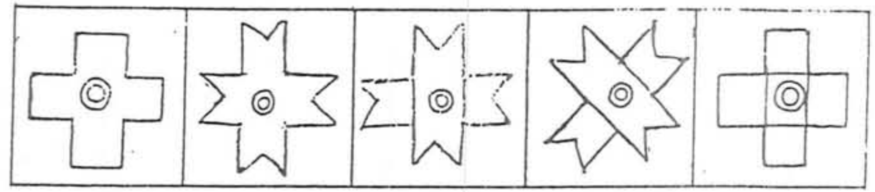




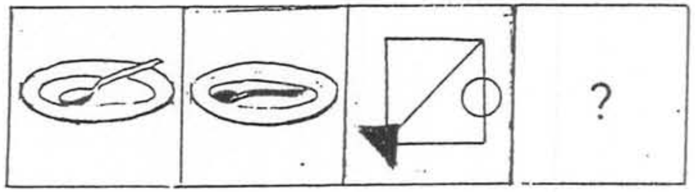
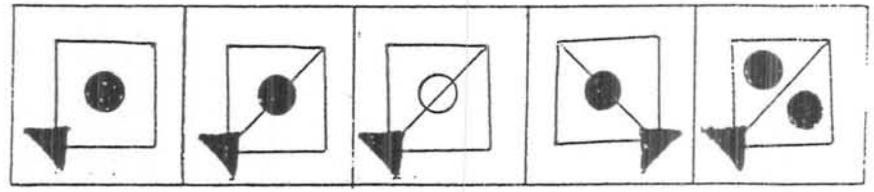
1 2 3 4 5

PROBLEM FIGURES سوالیہ اشکال

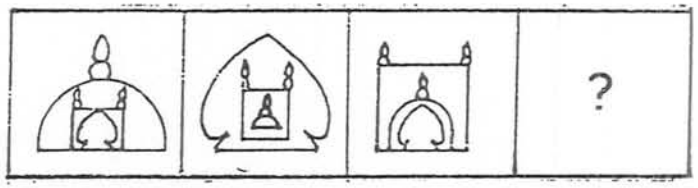
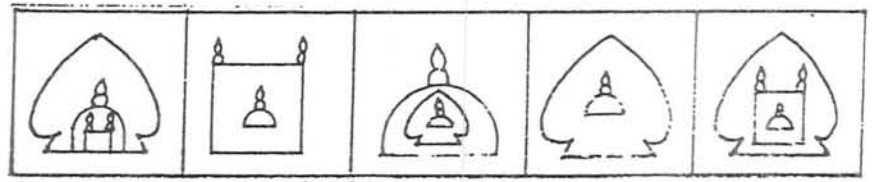
ANSWER FIGURES جوابی اشکال

11.  

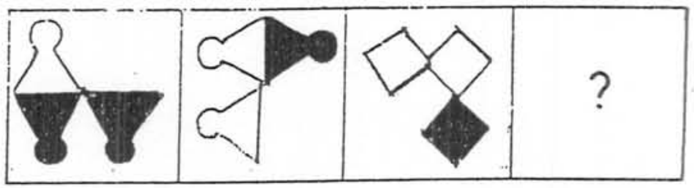
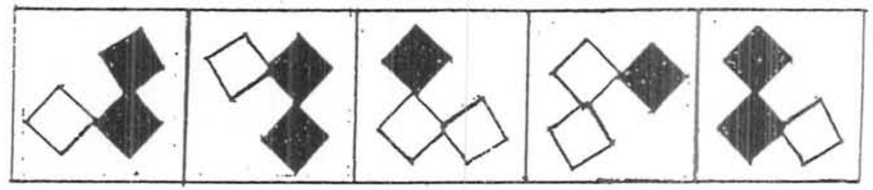
1 2 3 4 5

12.  

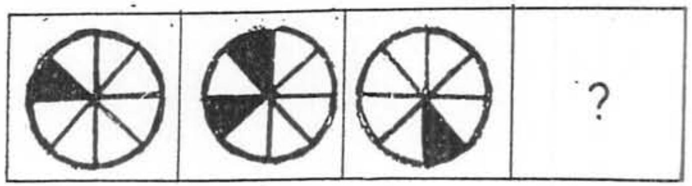
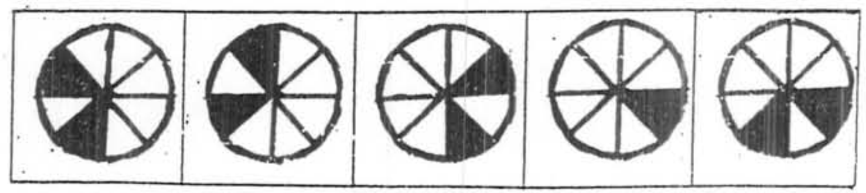
1 2 3 4 5

13.  

1 2 3 4 5

14.  

1 2 3 4 5

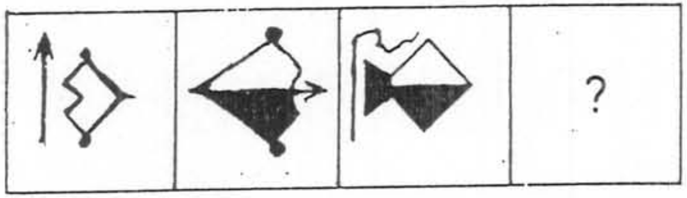
15.  

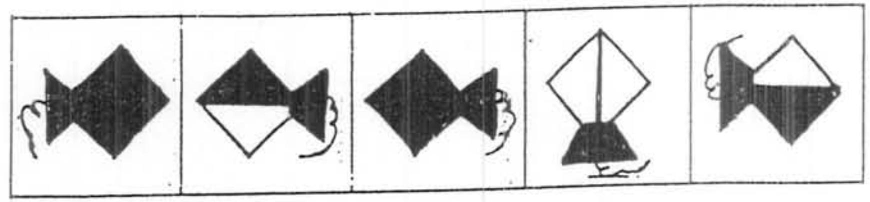
1 2 3 4 5

PROBLEM FIGURES سوالیہ اشکال

ANSWER FIGURES جوابی اشکال

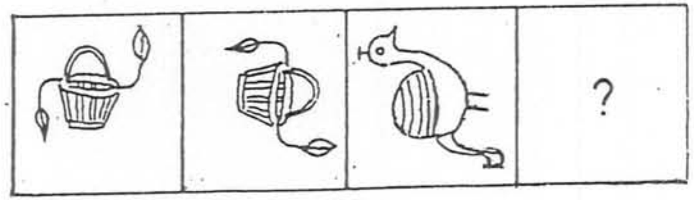
16

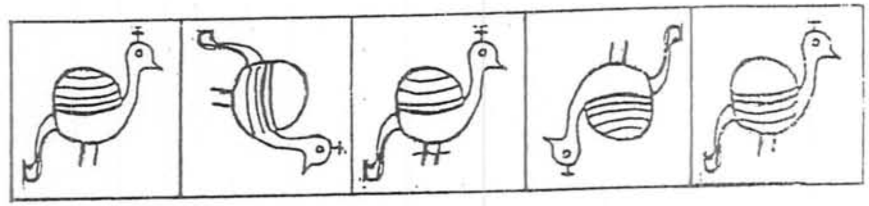




1 2 3 4 5

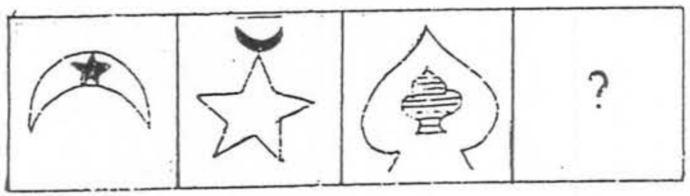
17.

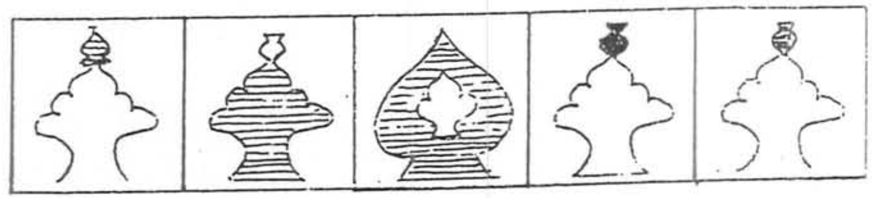




1 2 3 4 5

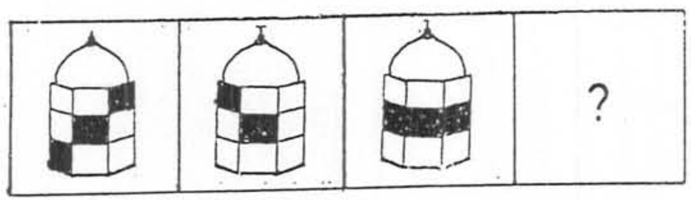
18.

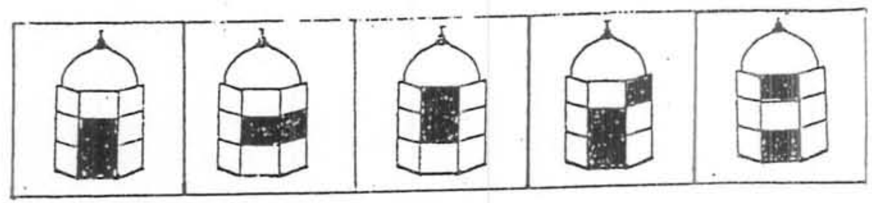




1 2 3 4 5

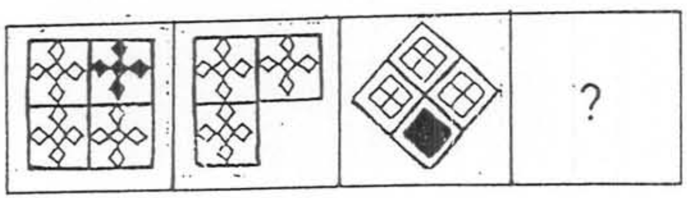
19.

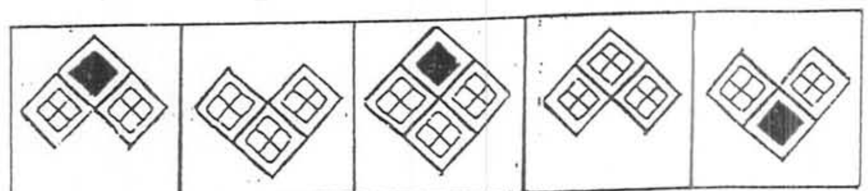




1 2 3 4 5

20.





1 2 3 4 5

ODD ONE OUT

Time limit – 7 Minutes

INSTRUCTIONS

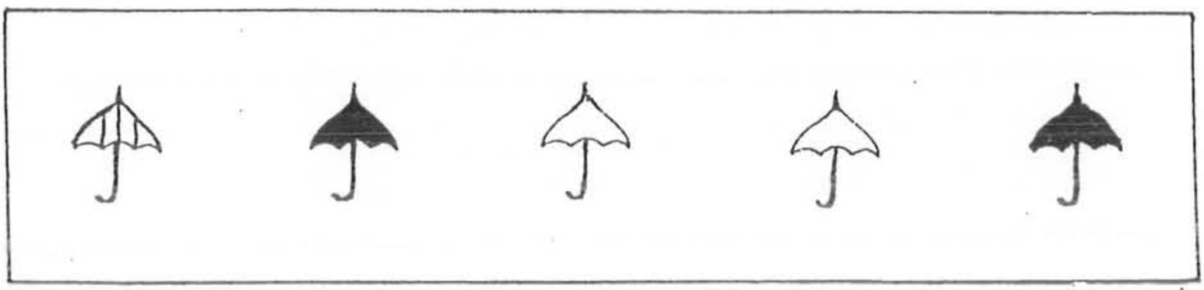
Each item of this test consists of five figures. One out of these five figures is different from the other four. Find out this figure and write its number on the answer sheet.

ہدایات

اس ٹیسٹ میں ہر سوال پانچ اشکال پر مشتمل ہے۔
ان پانچ اشکال میں ایک شکل باقی چار سے مختلف
ہے۔آپ نے اس شکل کو تلاش کرنا ہے اور اس
کے نیچے دیئے گئے نمبر کو جوابی فارم میں اس
سوال کے سامنے والے خانے میں لکھنا ہے۔

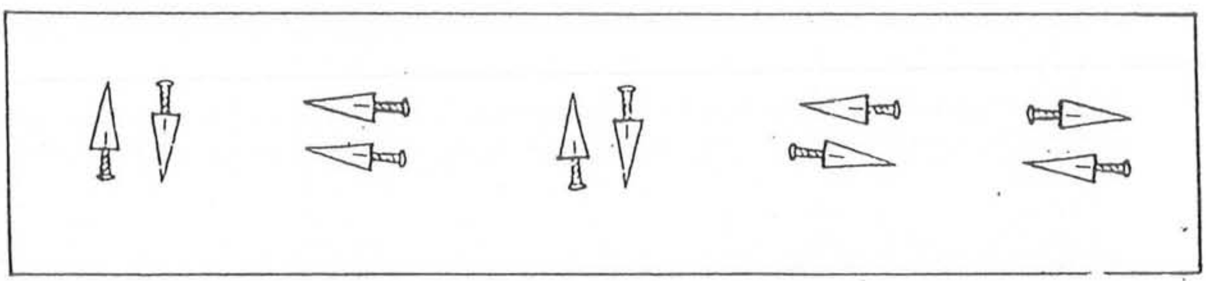
EXAMPLES

1.



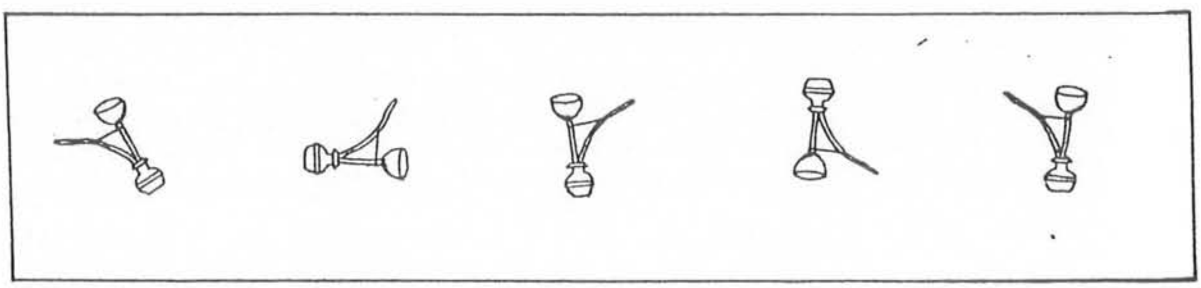
1 2 3 4 5

2.



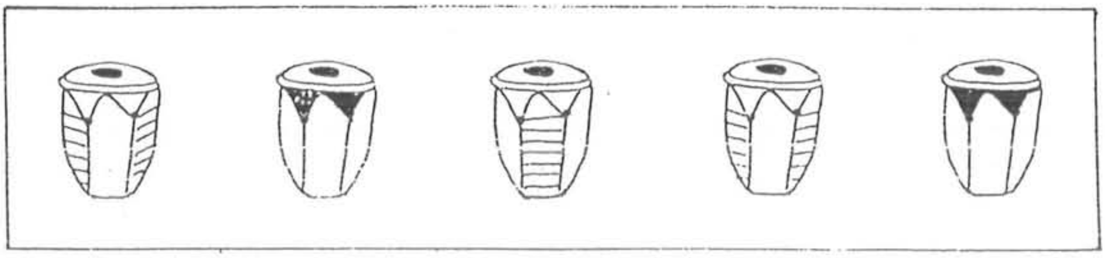
1 2 3 4 5

3.



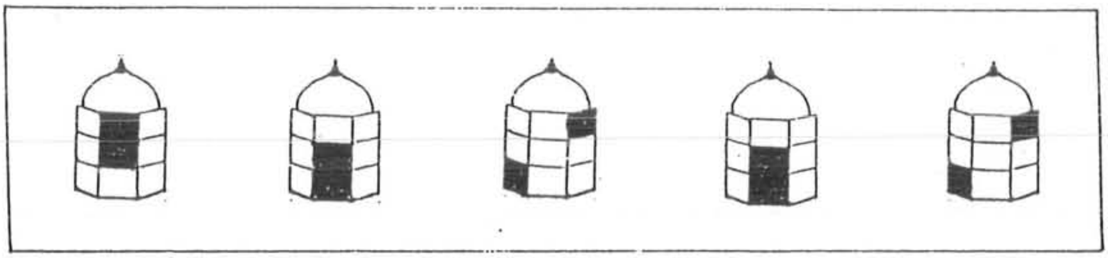
1 2 3 4 5

1.



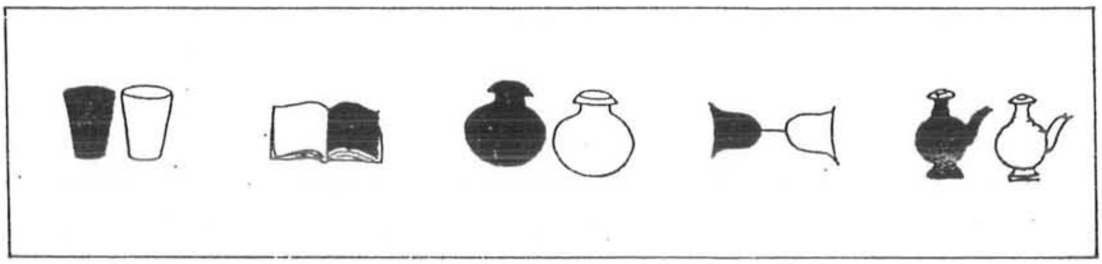
1 2 3 4 5

2.



1 2 3 4 5

3.

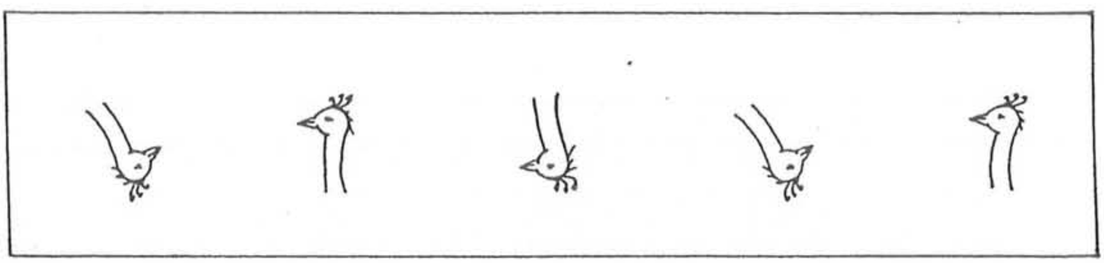


1 2 3 4 5

4.

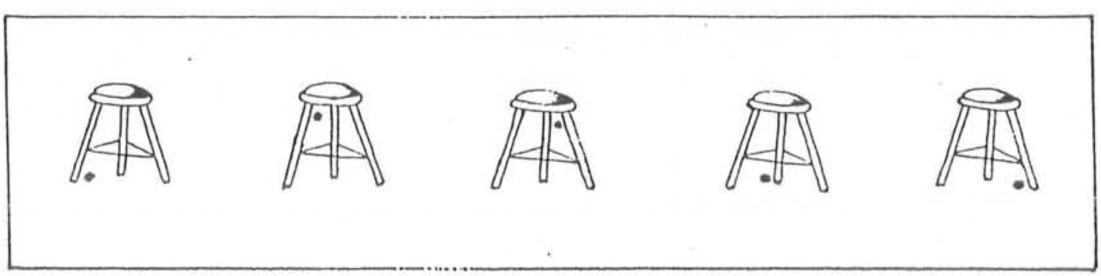
1 2 3 4 5

5.



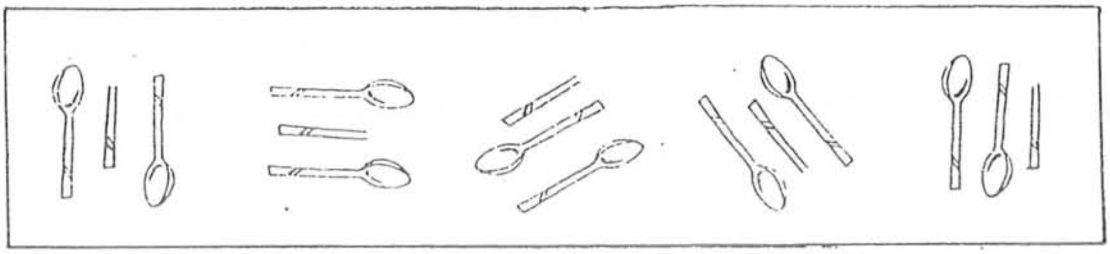
1 2 3 4 5

6.



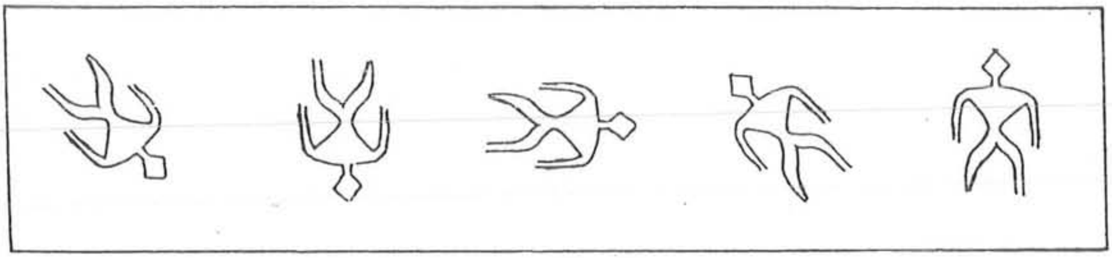
1 2 3 4 5

7.



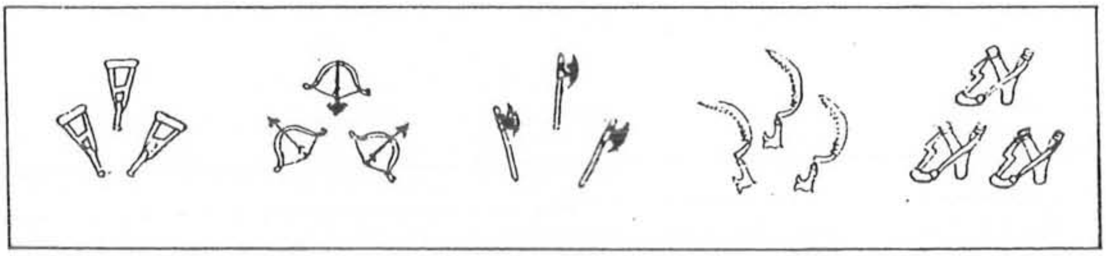
1 2 3 4 5

8.



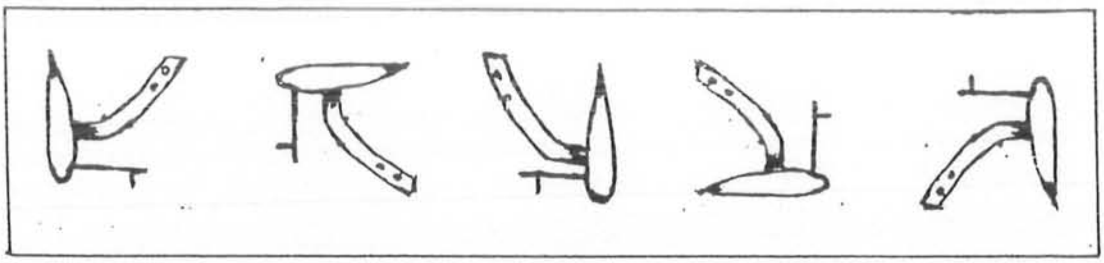
1 2 3 4 5

9.



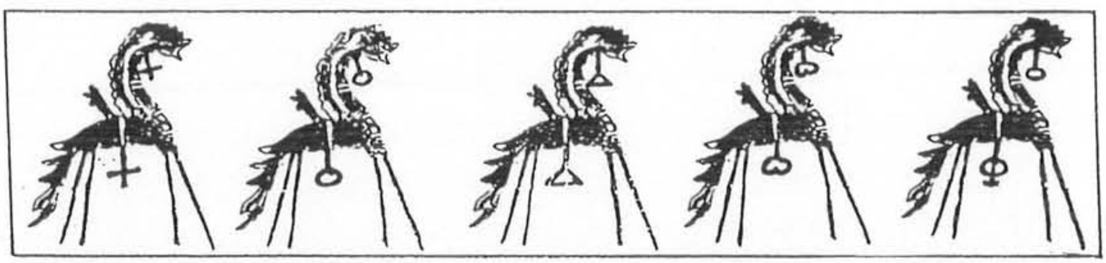
1 2 3 4 5

10.



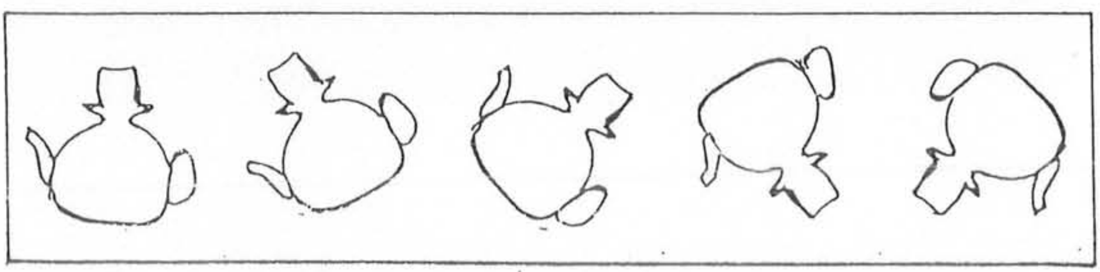
1 2 3 4 5

11.



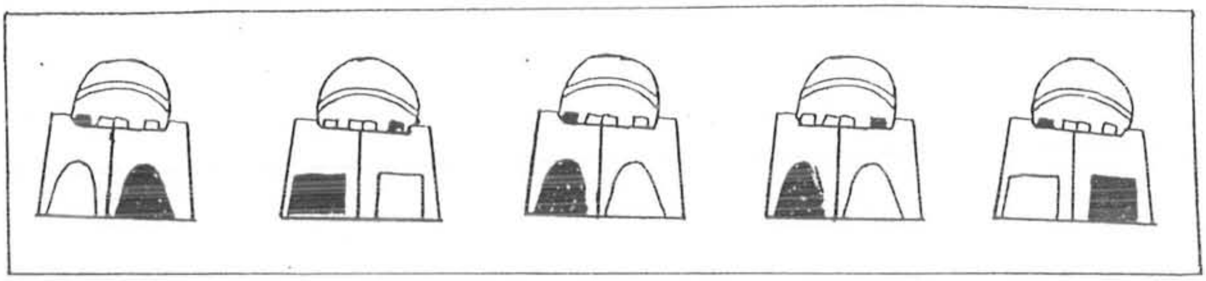
1 2 3 4 5

12.



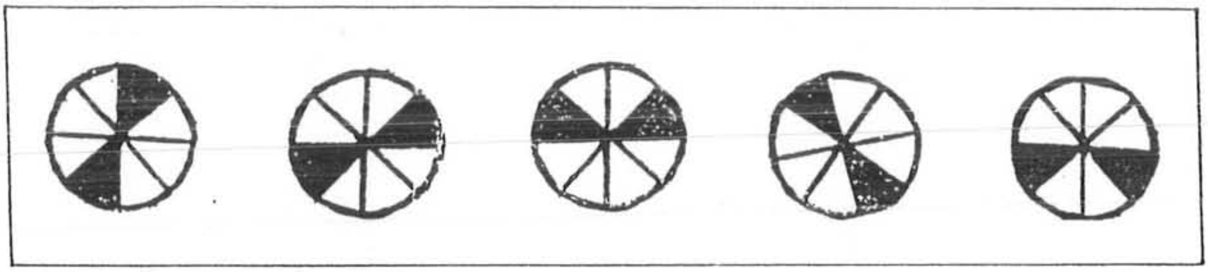
1 2 3 4 5

13.



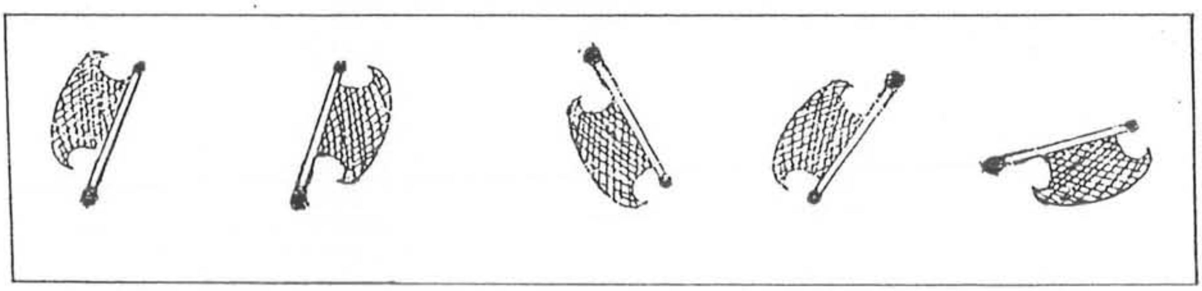
1 2 3 4 5

14.



1 2 3 4 5

15.



1 2 3 4 5

SIMILARITIES

Time limit – 8 Minutes

SIMILARITIES

INSTRUCTIONS

Each item of this test consists of two sets of figures. Figures on the left side are called probelm figures and those on the right side are called answer figures. There is a definite similarity between a problem figure / set of problem figures and one of the four sets of answer figures. Find out which set of the answer figures is most similar to the problem figures and write its number on the answer sheet.

ہدایات

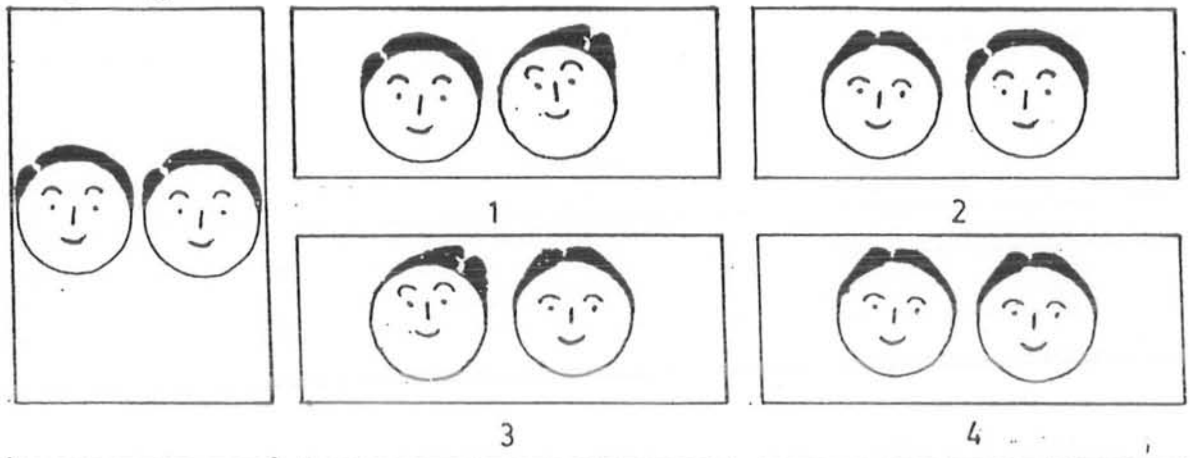
اس ٹیسٹ کے سوالات دو حصوں پر مشتمل ہیں۔ بائیں طرف سوالیہ اشکال اور دائیں طرف جوابی اشکال کے چار سیٹ ہیں۔ ہر سوالیہ شکل اور چار سیٹوں میں کسی ایک سیٹ کی اشکال کے درمیان ایک خاص تعلق یا مشابہت ہے۔ آپ نے اس سیٹ کا انتخاب کرنا ہے اور اس کے نیچے لکھے گئے نمبر کو جوابی فارم میں اس سوال کے سامنے والے خانے میں لکھنا ہے۔

EXAMPLES

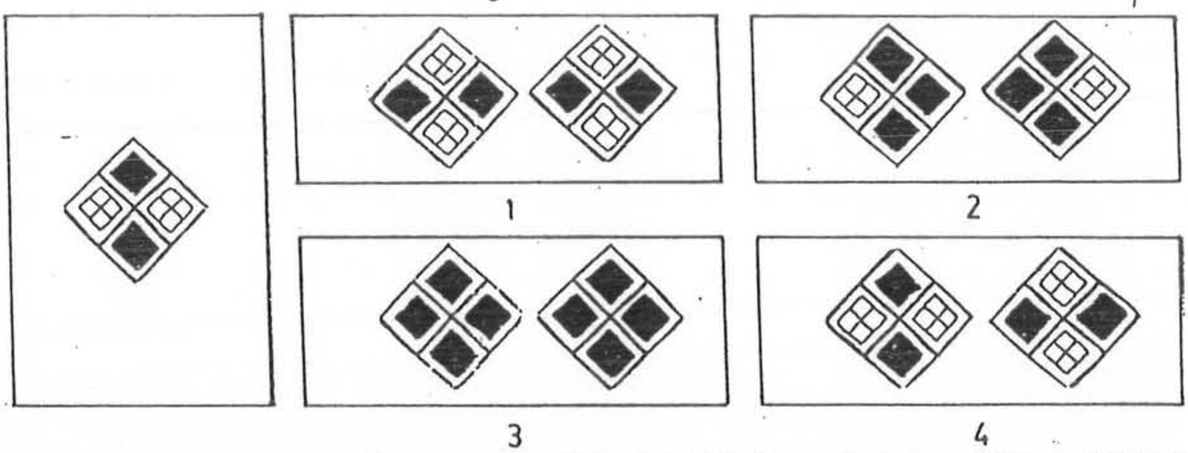
PROBLEM FIGURES سوالیہ اشکال

ANSWER FIGURES جوابی اشکال

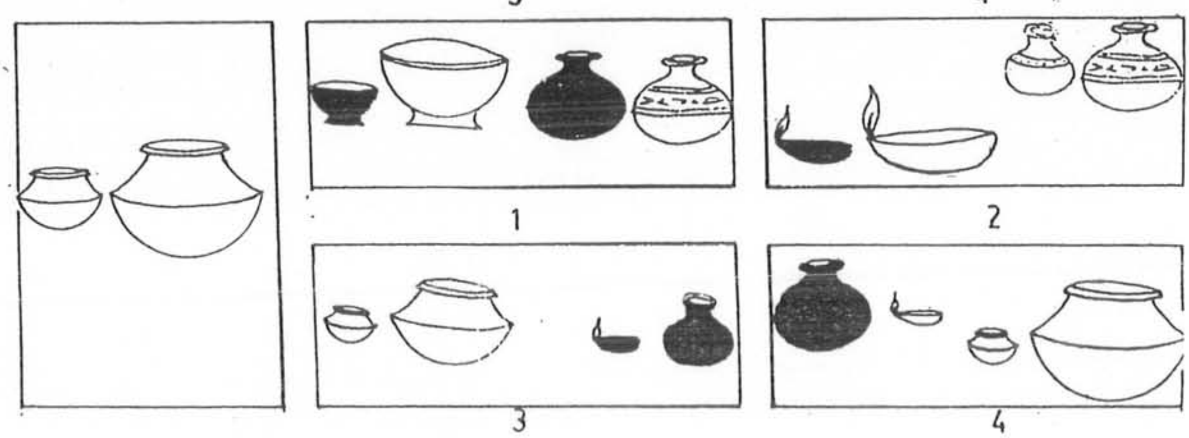
1.



2.



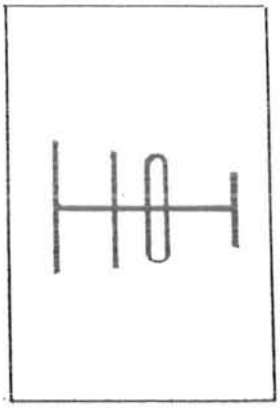
3.

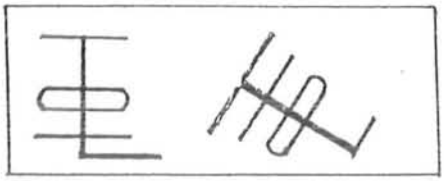


PROBLEM FIGURES

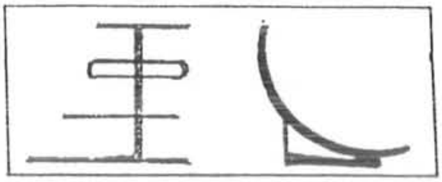
ANSWER FIGURES

1.

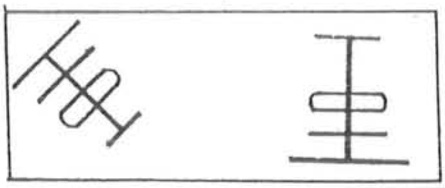




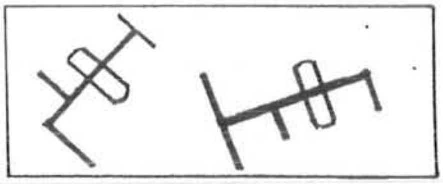
1



2

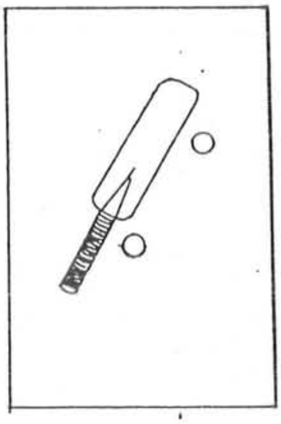


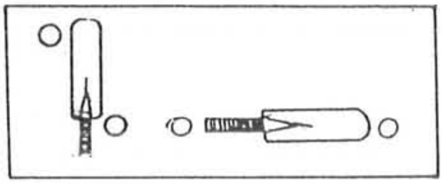
3



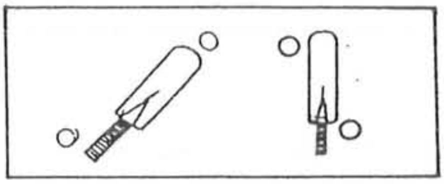
4

2.

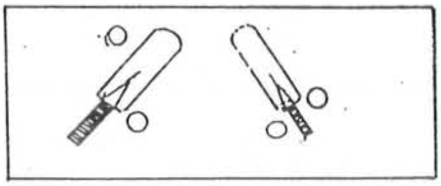




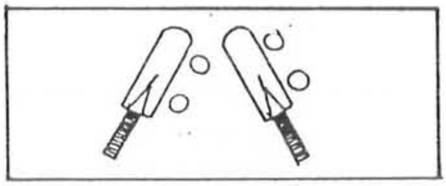
1



2

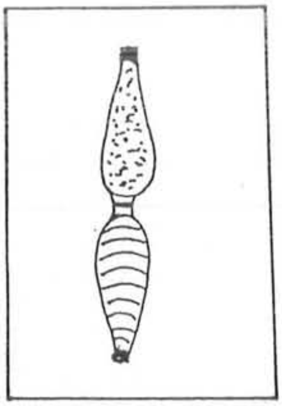


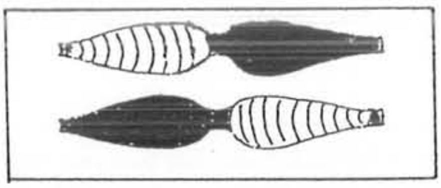
3



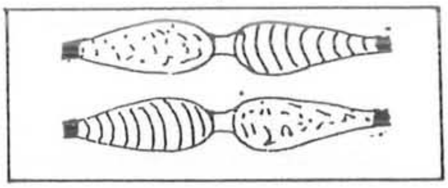
4

3.

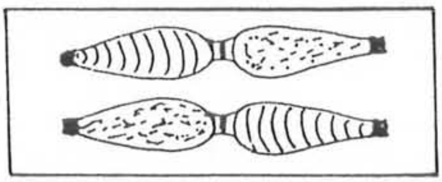




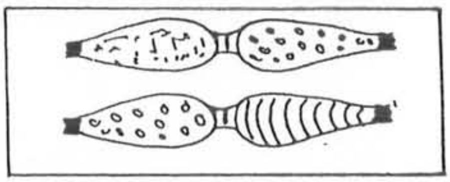
1



2

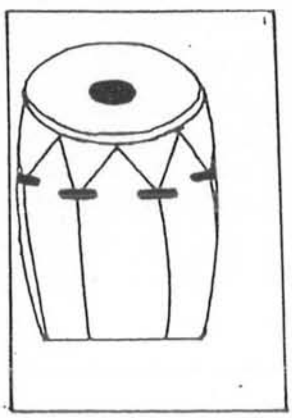


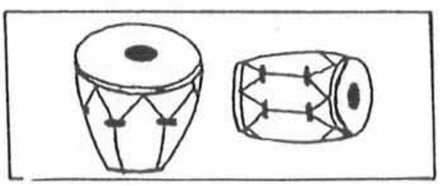
3



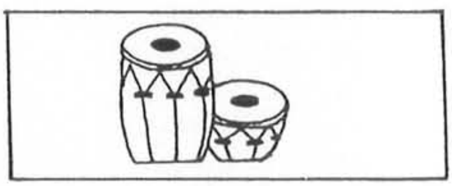
4

4.

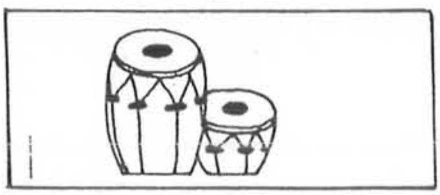




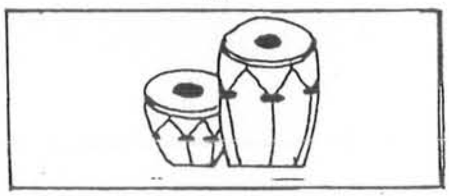
1



2



3

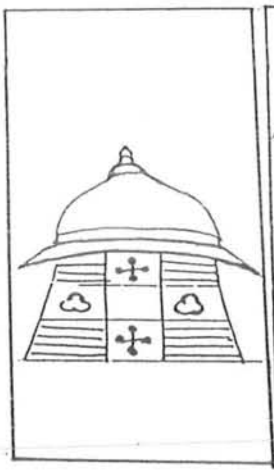


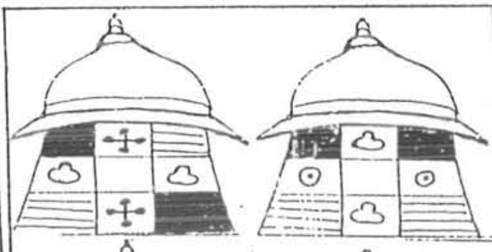
4

PROELEM FIGURES

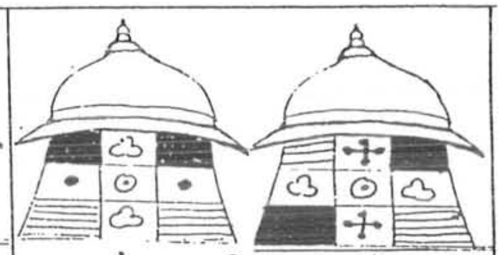
ANSWER FIGURES

5.

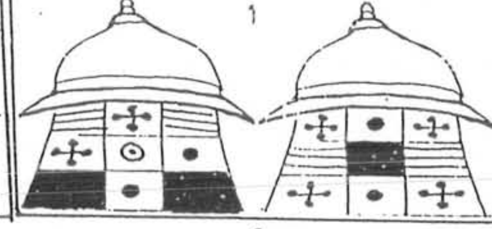




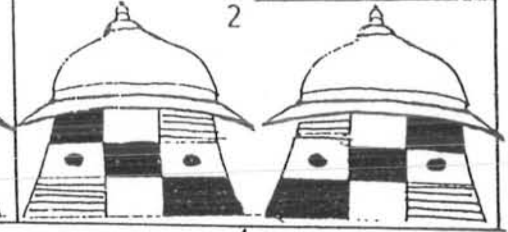
1



2

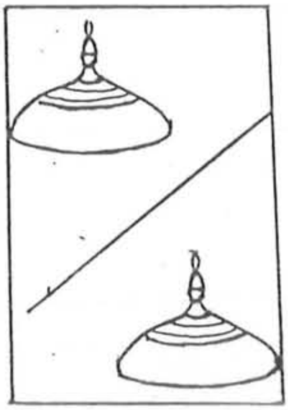


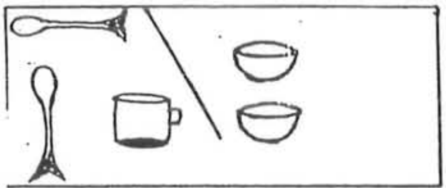
3



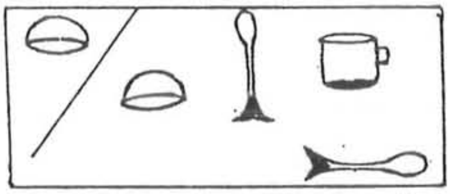
4

6.

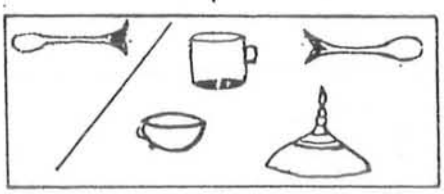




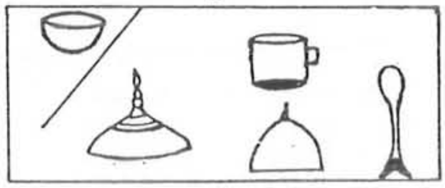
1



2

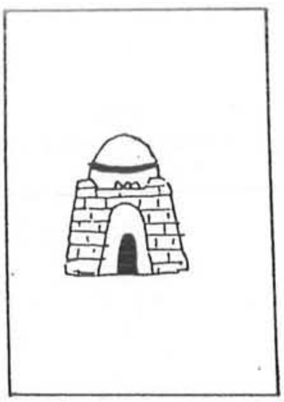


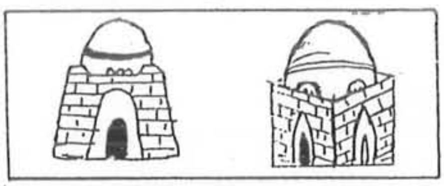
3



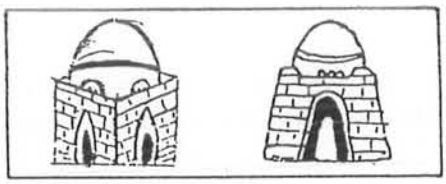
4

7.

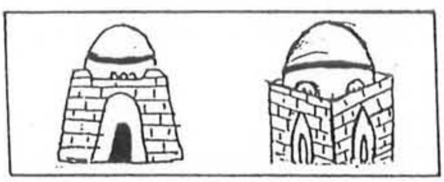




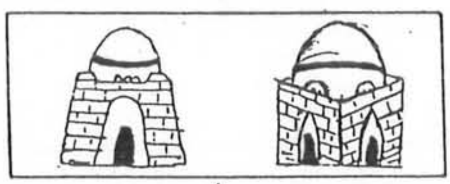
1



2

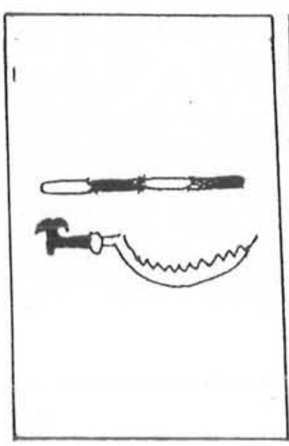


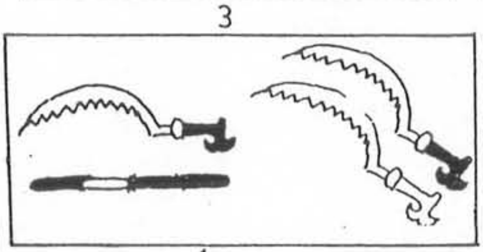
3



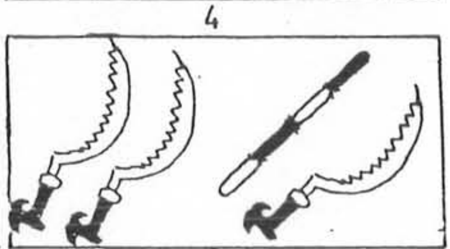
4

8.

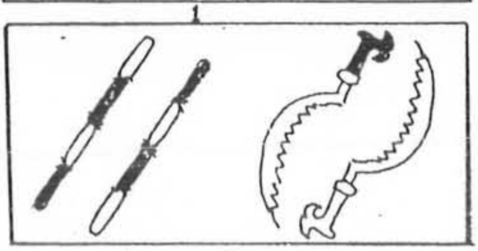




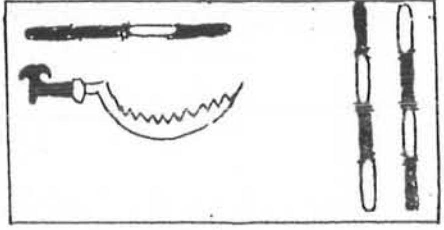
1



2



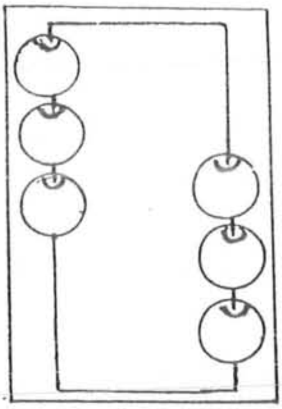
3

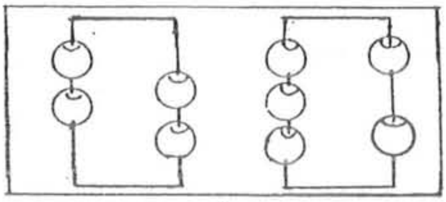


4

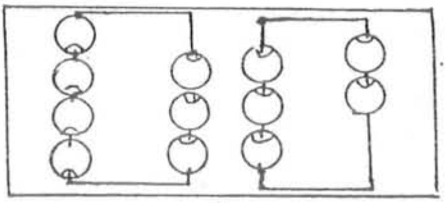
OBLEM FIGURES

ANSWER FIGURES

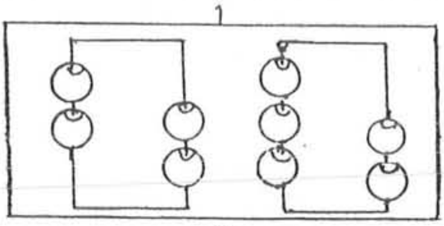




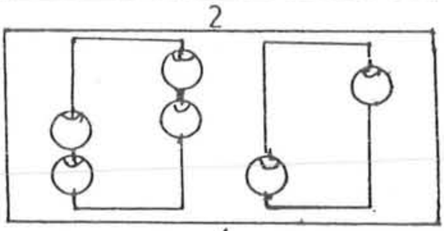
1



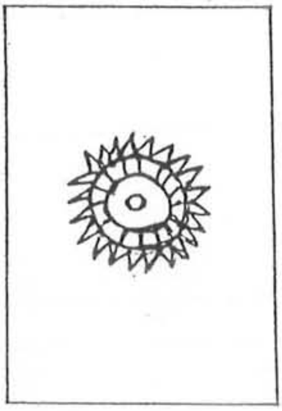
2

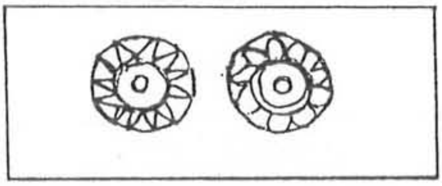


3

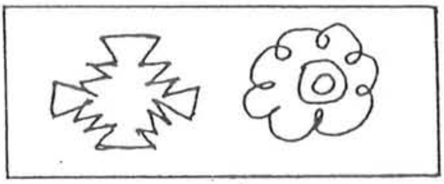


4

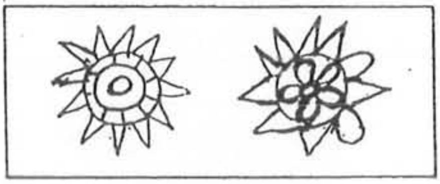




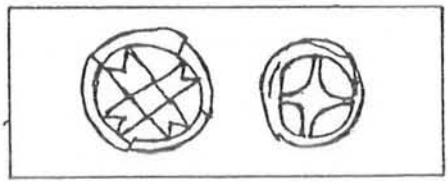
1



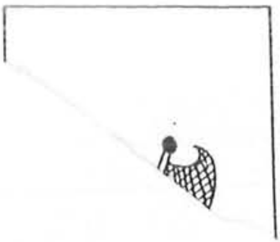
2

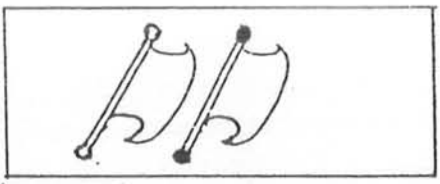


3

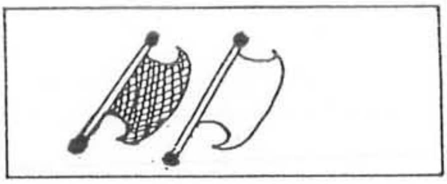


4

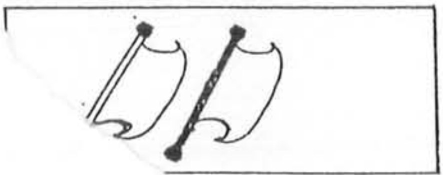


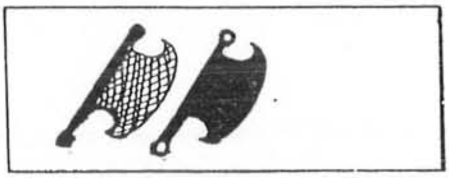


1

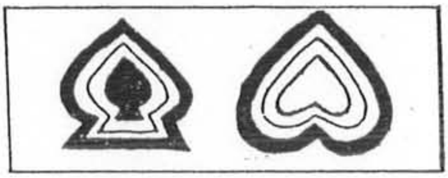


2

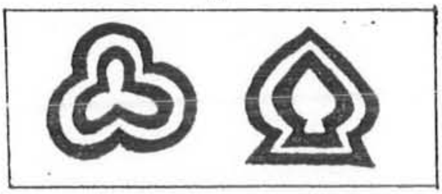




4



2

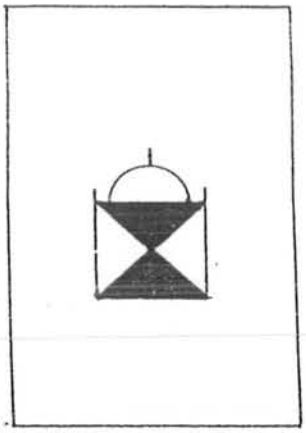


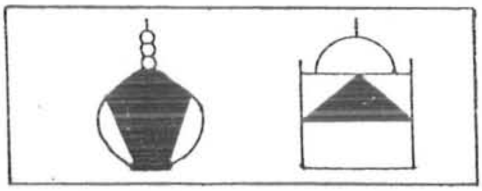
4

PROBLEM FIGURES

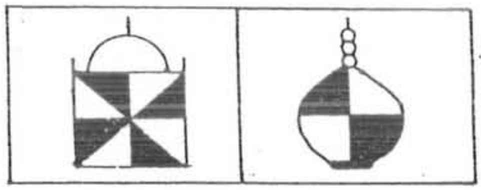
ANSWER FIGURES

13.

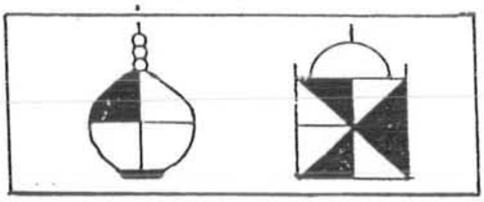




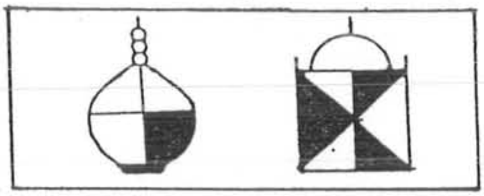
1



2

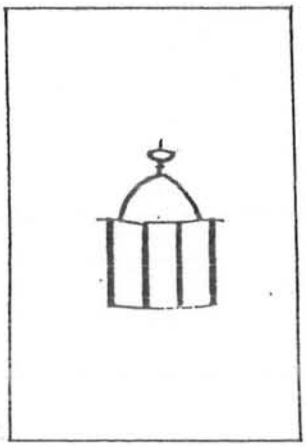


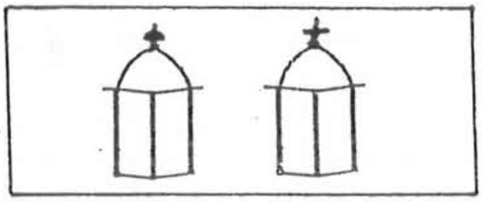
3



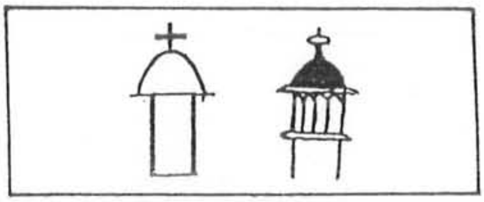
4

14.

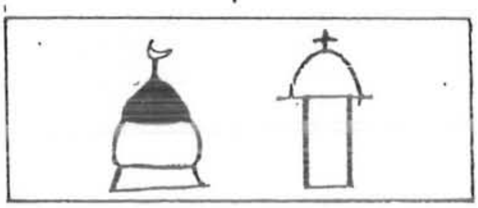




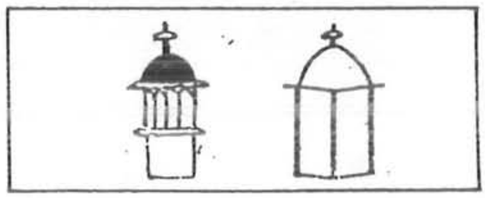
1



2

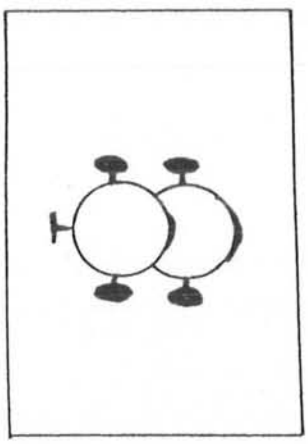


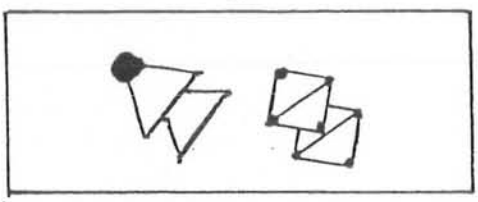
3



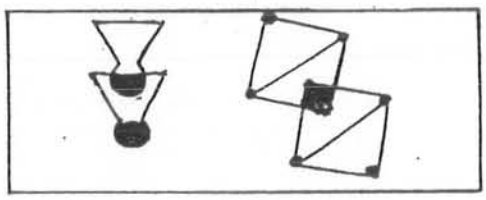
4

15.

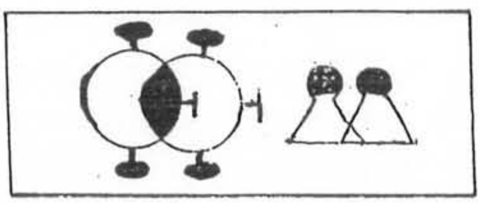




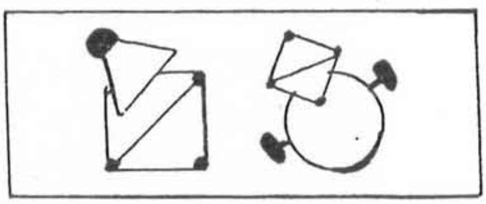
1



2



3



4