

# Development and Standardization of Intelligence Test for Children

Liaquat Hussain

Assistant Professor, Institute of Education and Research Gomal University DIKhan

Dr Asif Jamil

Associate professor, Institute of Education and Research Gomal University, DIKhan

E-mail: asifjamil72@hotmail.com

M. Junaid Siraji

Computer programmer, University WENSAM College DIKhan

Kiran Maroof

MEd, IER, Gomal University DIKhan, Khyber Pakhtunkhwa, Pakistan

Doi:10.5296/ijld.v2i5.2451

URL: <http://dx.doi.org/10.5296/ijld.v2i5.2451>

## ABSTRACT

*The problem under study was the development and standardization of intelligence test for children's. The main objectives of the study were to develop test items and Determining the Difficulty Level and Calculating the Discrimination Index of the test items. The study was significant because an intelligence test provides standardized ways of comparing a child performance with that of other children observed in the same situation (Anastasi, 1950). The study was delimited to only 600 children's of age group 6 to 11 years. Two statistical techniques difficulty level and discrimination index were used for item analysis.*

**KEY-WORDS:** Intelligence test, Standardization, Children

## INTRODUCTION

Different expert in the field of intelligence testing made their contribution in intelligence and its measurement. They proposed different factors on which intelligence is based. Different types of test are applied at different level. Unfortunately there is need for standardized intelligence test in our country. This study is aim to construct, try and standardize an intelligence test for children of age group 6-11 years. Hopefully, this study will provide a base for measurement of intelligence at primary / secondary school level in Pakistan.

Intelligence and its measurement remained a crucial issue in the later part of the nineteenth century. In order to construct an intelligence test the views of different psychologist and experts of the field have been taken into consideration. In the review of the related

literature various constructs of intelligence are described. Some of the important research studies regarding intelligence are reviewed in the literature review.

### **STATEMENT OF THE PROBLEM**

The problem under study was “the development and standardization of intelligence test for children’s”.

### **OBJECTIVE OF THE STUDY**

Following were the main objectives of the study.

1. Selection of intelligence test items for children of age group 6-11.
2. Determining the Difficulty Level of intelligence test items.
3. Calculating the Discrimination Index of the test items.

### **SIGNIFICANCE OF THE STUDY**

This study may be useful due to the following reasons.

1. An intelligence test provides standardized ways of comparing a child performance with that of other children observed in the same situation (Anastasi, 1950).
2. Intelligence test provides a profile of strength and weaknesses.
3. Intelligence tests are excellent predictors of scholastic achievement (Butcher, 1950).
4. Intelligence test assesses individual differences and provided useful reflection of cultural and biological differences among individual (Hunt, 1973).
5. Intelligence tests are valuable tools in working with handicapped children.

### **LIMITATIONS OF THE STUDY**

Following were the major limitation of the study.

1. This study cannot predict non academic intellectual activities.
2. This study cannot predict occupational success.
3. One of the limitations of this study is that it limits the concept of intelligence.
4. Intelligence testing requires as much more time, but no such time was available.
5. This study is only bound to find difficulty level and discrimination index of items in the intelligence test.

### **DELIMITATIONS OF THE STUDY:**

Following were the delimitation of the study.

1. The study was delimited only to school children.
2. The study was delimited to the children of age group 6 – 11 year.
3. The study was delimited to the sample of 600 children selected from rural and urban areas of District Dera Ismail Khan.

### **REVIEW OF RELATED LITERATURE**

#### **History of Intelligence Tests**

Interest in intelligence and in intelligence testing was an inherent part of the movement, beginning in the later part of the nineteenth century, that brought psychology into being as a separate discipline. Intelligence testing had its root in the field of general psychology and measurement. The physical method developed by E.H. Weber (1795-1878) and G.T.

Fechner (1801-1887), the study of difference limens by G.E. Muller (1850-1934) and F.M. Urban, and the statistical studies of higher mental processes initiated by Sir Francis Galton (1822-1911) formed the background for much of the work that would take place in the twentieth century.

Sir Francis Galton played an extremely influential role in the development of testing movement. The psychometric field was able to flourish because he originated two very important statistical concepts regression to the mean and correlation. Some regard Galton as the true father of the testing movement. In 1883 he published an important work "Inquiries into human faculty" in which he considered the problem involved in measuring mental characteristics.

Galton tried to measure sensory powers and thus intelligence. In 1882 he set up a testing booth in a London Museum and charged people a fee to have their learning, vision and other sensory motor equipment measured. One of Galton's students "Kari Pearson" worked out the basic equation of correlation coefficient that has proved so useful in the study of intelligence.

James Mckeen Cattle (1860-1944) established his own laboratory at the University of Pennsylvania. In an article published in 1890 in *Mind* he used the term "Mental test". So different measures were described, most of them of a sensory and motor nature and differing little from those of Galton.

Cattell stressed that psychology must rest on a foundation of measurement and experimentation. He foresaw the practical application of test as tools for the selection of people for training and for diagnostic evaluation. Clark Wissler (1901) was another investigator who sought to determine the validity of some of the test that was thought to be related to be cognitive processes.

At about the same time in France, Alfred Binet (1857 – 1911), Victor Henri (1872-1940) and Theodore Simon (1873-1961) were developing methods for the study of a variety of mental functions. These investigators found the key to the measurement of intelligence by focusing on higher mental processes instead of on simple sensory function. Their work appears in the form of Binet-Simon Scale (1905). The 1905 scale might be considered the first modern intelligence test of mental ability. With the introduction of the Binet-Simon scales, the testing movement began to flourish in the United States. Lewis Terman standardized the Binet-Simon scale in 1916, Robert Yerkes and his colleagues however found fault with the age-scale format of the Binet-Simon scales and developed their own point scale.

Many other specialized tests were also developed to evaluate specific facets of cognitive ability. Testing in school, clinics, industry and the military became a common practice and influenced public affairs, business, and scientific psychology.

In 1939, David Wechsler introduced the Wechsler-Bellevue intelligence scale in designing this scale Wechsler studied the standardized test that were available during the later 1930s, and selected eleven different subsets to form the scale. Source for the subsets included the Army Alpha, Stanford-Binet test, Healy picture completion tests, army group examination, Kohs-Block design test and Army beta.

Wechsler focus on the global nature of intelligence. Wechsler considered intelligence to be a part of the large whole of personality itself. The Wechsler designed many scales of intelligence including Wise, Wisc-R, Wais-R, WPPSI etc. The overall IQ obtained from the scale represented an index of general mental ability.

### **Testing of Intelligence**

Intelligence testing has generated much controversy in recent year. Some writer maintains that the development of intelligence tests is one of the most significant contributions

in the field of psychology, while others believe that intelligence tests have many serious shortcomings.

Our intelligence tests which measure success in school quite effectively are value laden. They represent primarily such social values as schooling, verbal abilities, abstraction and concept formation skills. These are important skills in western societies but they are not the only ones.

Attempts are being made to find ways to measure and reward different kind of skills (e.g creativity) from those currently associated with intelligence tests. Such efforts are needed because our tests tap only part of the spectrum of human abilities. Still our current intelligence test does their job well. They predict success in school, and they measure some of the important skills needed in our industrialized society.

### **Assessment of Intelligence**

No other are of assessment has generated as much attention, controversy, and debate as the testing of what we call “intelligence”. For centuries philosophers, Psychologists, educators, and lay people have the meaning of intelligence. Numerous definitions of the term intelligence have been proposed, with each definition serving as a stimulus for counter definitions and counter proposals. Several theories have been advanced to describe and explain intelligence and its development such as Cattle-Horn, Gardner, Kaufman, Sternberg, and Guilford (Flanagan, Genshaft & Harrison, 1997). The extent to which the intelligence is genetically or environmentally, determined has been of special concern, Genetic determinists, environmental determinists, and inter actionists have all observed difference in the intelligence test performances of different populations of children.

Both the interpretations of group difference in intelligence measurements and the practice of testing the intelligence of school children have been topics recurrent controversy and debate, aired in professional journals, in the popular press, and on television. In some instance the courts have acted to curtail or halt intelligence assessment in the public school, in other, the courts have defined what composes intelligence assessment. Debate and controversy have flourished about whether intelligence test should be given, what they measure, and how different level of performance attained by different populations is to be explained.

No one, however, has seen a specific thing called “intelligence”. Rather, we observe differences in the ways people behave-either differences in everyday behavior in a variety of situations or differences in response to standard stimuli or sets of stimuli, then we attribute those differences to something termed intelligence.

In this sense intelligence is an inferred entity, a term or construct we use to explain differences in present behavior and to predict differences in future behavior.

We have repeatedly stressed the fact that all tests, including, therefore, intelligence tests assess sample of behavior, Regardless, of how an individual’s performance on any given tests is viewed and interpreted, intelligence tests- and on the item on those tests-simply sample behaviors. A variety of different kinds of behavior samplings are used to assess intelligence, in most cases, the kind of behavior sample reflect a test author’s conception of intelligence. The behavior sample are combined in different ways in which they view the concept of intelligence, here we review the kinds of behaviors sampled by intelligence tests, with particular emphasis on the psychological demands of different test items as a function of pupil characteristics.

### **Intelligence tests as sample of behavior**

There is a hypothetical domain of items that could be used to assess intelligence. In practice, it is impossible to administer every item in the domain to a student whose intelligence we want to assess. The characterization of behaviors sampled by intelligence tests is complex. Some persons have argued, for example, that intelligence tests assess a student’s capacity to

profit from instruction, whereas others, argue that such tests assess merely what has been learned, some have characterized intelligence tests as either verbal or nonverbal, some characterized intelligence tests as either culturally biased or culturally fair. In actuality, nearly any contention regarding what intelligence tests measure can be supported. The relative merit of competing options, theories and contentions is primarily a function of the interaction between the characteristics of an individual and the psychological demands of items in an intelligence test. It is also a function of the stimulus and response requirements of the items.

Similar behavior may assess in different ways. In assessing vocabulary, for example, the examiner may ask the pupils to define words, to name pictures, to select a synonym of a stimulus word or to point to pictures depicting words read by the examiner. All found kinds of assessments are called “vocabulary tests” yet they sample different behaviors. The psychological demands of the items change with the ways the behavior is assessed.

In evaluating the performance of individuals on intelligence tests, teachers, administrators, counselors, and diagnostic specialists must go beyond test names and scores to look at the kinds of behaviors sampled on the test. They must be willing to question the ways test stimuli are presented to question the response requirements, and to evaluate the psychological demands placed on the individual.

### **Behavior Sampled by Intelligence Tests**

Regardless of the interpretation of measured intelligence, it is a fact that intelligence tests simply sample behaviors. This section describes the kinds of behaviors sampled, including discrimination, generalization, motor behavior, general knowledge, induction, comprehension, sequencing, detail recognition, analogical reasoning, pattern completion, abstract reasoning, and memory.

#### **Discrimination**

Intelligence test-items that sample skill discrimination usually, present a variety of stimulus, and ask the student to find the one that differs from all the others. Figural, symbolic, or semantic discrimination may be assessed. Figure 2.1 illustrates items assessing discrimination: both, items ‘a’ and ‘b’ assess semantic discrimination. In each case, the student must identify the item that differs from the others. The psychological demand of the items differs, however, depending on the student’s age and particular set of background experiences.

Figure 2.1 Items that assess symbolic and semantic Discrimination

#### **Symbolic Discrimination**

- |    |   |   |   |   |
|----|---|---|---|---|
| a. | 4 | A | Q | W |
| b. | a | 3 | 5 | 8 |

#### **Semantic Discrimination**

- |    |          |        |         |          |
|----|----------|--------|---------|----------|
| c. | elephant | horse  | monkey  | truck    |
| d. | Hispanic | French | Arabian | Germanic |

#### **Generalization**

Items assessing generalization present a stimulus and ask the student to identify which of several response possibilities goes with the stimulus. Again, the content of the item may be figural, symbolic, or semantic, the difficulty may range from simple matching to a more difficult type of classification.

### **Motor Behavior**

Many items on intelligence tests require a motor response. The intellectual level of very young children, for example, is often assessed by items requiring them to throw objects, walk, follow moving objects with their eyes, demonstrate a pincer grasp in picking up objects, build block towers, and place geometric forms in a recessed-form board. Most motor items of higher age levels are actually visual motor items. The student may be required to copy geometric designs, trace, paths through a maze, or reconstruct design from memory. Obviously, because motor responses can be required for items assessing understanding and conceptualization, many items assess motor behavior at the same time that they assess others behaviors.

Items that assess Symbolic and Semantic Generalization

### **Symbolic Generalization**

- c. J        H        8        6        9  
d. 81       21       23       26       25

### **Semantic Generalization**

- e. tree       car       man       horse       walk  
f. salvia       flashlight       frog       tulip       banana

### **General Knowledge**

Items on intelligence tests sometimes require a student to answer specific factual questions, such as “in what direction would you travel if you were go to from Poland to Argentina?” “And” what is the cube root of 8?” Essentially, such items are like the kinds of the items in achievement tests, they assess primarily what has been learned.

### **Vocabulary**

Much different kind of test items are used to assess vocabulary. In some cases, the student must name picture, and in others, she or he must point to object to response to words read by the examiner. Some vocabulary items require the student to product oral definitions of words, whereas others call for reading a definition and selecting one of several words to match the definition. Some test score a students definitions of words as simply pass and fail, others uses a weighted scoring system to reflect the degree of abstraction used in defining words. The Wechsler intelligence scale for Children-III, for example, assigns 0 points to incorrect definitions. I point to definitions that are descriptive (an orange is round) or functional (an orange is to eat) and 2 points to more abstract definitions (an orange is a citrus fruit).

### **Induction**

Induction item present a series of example and require the student to induce a governing principle. For example, the student is given a magnet a several different cloth, wooden, and metal objects and is asked to try to pick up the objects with the magnet. After

several trials, the student is asked to state a rule or principle about the kinds of objects that magnets can pick up.

### **Comprehension**

There are three kinds of items used to assess comprehension, items related to directions, to printed material and to societal customs and mores. In some instance, the examiner presents a specific situation and ask what actions the student would take (for example, “what would you do if you saw a train approaching a washed-out bridge?”). In other cases, the examiner reads the paragraphs to a student and then asks specific questions about the content of paragraphs. In still other instance, the student is asked questions about social mores, such as, “why should we keep promises?”

### **Sequencing**

Items assessing sequencing consists of series of stimuli that have a progressive relationship among them. The student must identify a response that continues the relationship.

### **Detail Recognition**

In general, not many tests o test items assess detail recognition. Those that do evaluate the completeness and detail with which a student solves problems. Certain drawing tests. Such as the Good enough-Harris Draw- A-Person Test, evaluate a student’s drawing of a person on the basis of inclusion of detail. The more details in a student’s drawing, the more credit the student earns. In other instances, items require a student to count the blocks in pictured piles of blocks in which some of the blocks are not directly visible, to copy geometric designs, or to identify missing part in pictures. To do so correctly, the student must attend to detail in the stimulus drawings and must reflect this attention to detail in making responses.

### **Analogical Reasoning**

“A is B as C is to ..... “ is the usual form for analogies. Element A is related to element B. The student must identify the response having the same relationship to element C as B to A.

### **Pattern completion**

Some test and test items require a student to select form several possibilities the missing part of a pattern or matrix.

### **Abstract Reasoning**

A variety of items on intelligence tests sample abstract reasoning, ability. The standford-Binet, intelligence scale, for example, presents absurd verbal statements and pictures and ask the student to identify the absurdity. It also includes a series of proverbs, the essential meaning of which the student must state. In the Stanford-Binet and other scales, arithmetic-reasoning problems are often though to assess abstract reasoning.

Figure 2.2 analog items

- a. man: boy: : woman -----girl mother daughter aunt
- b. tapeworm: plat helminthes : starfish: -----echinoderm mollusca water porifera
- c. variance : standard deviation : : 25 -----4 5 625 747

### **Coping with Dilemmas in Current Practice**

The practice of assessing children intelligence is currently marked by controversy. However, much of that controversy could be set aside if intelligence tests were viewed appropriately. Intelligence tests imply assess sample of behavior, and different intelligence test sample different behaviors. For that reason, it is wrong to speak of a person’s IQ. Instead, we can refer only to a person’s IQ on specific test. An IQ on the Stanford-Binet IV intelligence scale is not derived from the same samples of behaviors as an IQ on any other intelligence test.

Because the behavior samples are different for different tests, educators, and others must always ask, “IQ on what test?”. It is helpful to understand that for the most part, the particular kinds of items or subtests found on an intelligence test are a matter of the way in which a test author defines intelligence and think about the kinds of item and subtests that assess it.

The same test may make different psychological demands on various test takers, Depending on their ages and acculturation. Test result mean different things for different students. It is imperative that we be especially aware of the relationship between a person’s acculturation and acculturation of the norm group with which that person is compared.

### **Memory**

Several different kinds of tasks assess memory, repetition of sequences of digits presented orally, reproduction of geometric designs from memory, verbatim repetition of sentences, and reconstruction of the essential meaning of paragraphs or stories. Simply saying that an item assess memory is too simplistic. We need to ask, memory for what?. The psychological demand of memory task changes in relation to both the method of assessment and the meaningfulness of the material to be recalled.

## **METHODOLOGY**

### *Population*

All the children’s of age group 6-11 years comprised the population of the study.

### *Sample*

A sample of 600 children was selected from urban and rural areas as under: -

- |                                       |   |    |
|---------------------------------------|---|----|
| 1. Male urban children of age 6 years | = | 25 |
| 2. -----7-----                        | = | 25 |
| 3. -----8-----                        | = | 25 |
| 4. -----9-----                        | = | 25 |
| 5. -----10-----                       | = | 25 |
| 6. -----11-----                       | = | 25 |
| 7. Male Rural -----6-----             | = | 25 |
| 8. -----7-----                        | = | 25 |
| 9. -----8-----                        | = | 25 |
| 10. -----9-----                       | = | 25 |
| 11. -----10-----                      | = | 25 |
| 12. -----11-----                      | = | 25 |
| 13. Female urban -----6 Years         | = | 25 |
| 14. -----7-----                       | = | 25 |
| 15. -----8-----                       | = | 25 |
| 16. -----9-----                       | = | 25 |
| 17. -----10-----                      | = | 25 |
| 18. -----11-----                      | = | 25 |
| 19. Female Rural-----6-----           | = | 25 |
| 20. -----7-----                       | = | 25 |
| 21. -----8-----                       | = | 25 |
| 22. -----9-----                       | = | 25 |
| 23. -----10-----                      | = | 25 |



24. -----11----- = 25

### **Procedure**

Criteria for selection of test items

The test item must satisfy the following criteria.

1. An effort was made to select those items which were considered interesting according to judgment of the researcher in order to reduce unwanted variation in performance resulting from poor motivation or flagging attention.
2. Socially and culturally neutral items were selected so that they neither favor an individual nor put him at a disadvantage because of the particular group to which he/she belongs.
3. Those items were selected which were considered neither very easy nor very difficult.

### **Administration of the Test**

After the selection of test items different sub-tests were arranged according to the age level of the students. As describe in the sample the 100 photo copies of each age level were prepared and these were tested on trial bases to the students of urban / rural and male / female schools. Instructions were given to the children for each part of a test. Each subtest was assigned specific time in which children were required to complete the subtest and a watch was used for this purpose. Test administration procedure was similar for all the participants.

Tests were marked using a standard procedure in which score of +1 was given for each item passed. Thus total 50 marks were assigned for each test.

### **Data Analysis**

After collecting the data, it was arranged in tabular form and two statistical techniques difficulty level and discrimination index were used for item analysis.

To find difficulty level and discrimination index the test score were divided into three groups, the highest 27 Percent, the middle 46 Percent, and the lowest 27 Percent. The difficulty level was calculated with the help of the formula  $(a + b) / 2$  where 'a' represent the percentage of the students doing the item right in the high achievers group and 'b' stands for the percentage of the students doing the item right in the low achievers group. The discrimination index was determined by the difference between the percentages of the students doing the item right in the high achievers and low achievers group i.e., discrimination index = a – b.

### **PRESENTATION AND ANALYSIS OF DATA**

Although there are several different process of item analysis giving indices of quality, traditionally two have been used, the difficulty level and the discrimination index. These are quite straight forward item statistics that can be used by class room teachers for analyzing their tests as well as being used for more extensive testing situations. The difficulty level is related to the difficulty of the item in terms of percentage of the students doing the item right. The large the difficulty level the easier the item, the smaller the difficulty level the more difficult the item. For example if an item has a difficulty level of 0.90 this means that 90 Percent of the persons taking the tests answered that item correctly, it clearly is a very easy test items.

Discrimination of the test item means that the test item is effective in separating those with high scores on the total test from those with low total test scores. It seems reasonable that

if an item is a good discriminator, students with high test scores will tend to get it correct and those with low test scores will respond incorrectly.

The two difficulty levels and discrimination index the test score were divided into three groups, the highest 27 percent, the middle 46 percent and the lowest 27 percent. The difficulty level was calculated with the help of the formula  $(a + b) / 2$  where 'a' represent the percentage of the students doing the item right in the high achievers group and 'b' stands for the percentage of the students doing the item right in the low achievers group. The discrimination index was determined by the difference between the percentages of the students doing the item right in the high achievers and low achievers group i.e., discrimination index =  $a - b$ .

The two properties of an item namely its difficulty level and its power to discriminate between good and poor students form the basis for judging its adequacy and worth. Item with extreme values of difficulty level and discrimination indices were not found acceptable for inclusion in the item pool, because too easy items did not have the capability to discriminate between students and too difficult items also tended to be low in discrimination value.

An item with 50% difficulty level is considered to be an ideal test item. However research shows that items with discrimination indices ranging from 16% to 84% could be included preferably. I also followed this criterion. However some expert of the field such as Ebel and Frisbie (1986, P.234) also accept it as valid beyond this range. But in no case items with discrimination indices less than or equal to zero were accepted.

**TABLE1: SUMMARY OF DIFFICULTY LEVEL OF ITEMS**

Range of difficulty level	Age group 6	Age group 7	Age group 8	Age group 9	Age group 10	Age group 11	Total
Below 16%	0	0	2	1	2	0	5
16 to 40%	1	0	1	6	4	3	15
41 to 60%	4	1	2	7	8	14	36
61 to 84%	23	20	36	32	23	17	151
Above 84%	22	29	9	4	13	16	93
Total	50	50	50	50	50	50	300

It is evident from the above table that the number of items that fall out of range 16% to 84% difficulty level are 22, 29, 11, 5, 15, and 16 respectively in tests of age level 6,7,8,9,10 and 11.

**TABLE2: SUMMARY OF DISCRIMINATION INDEX OF ITEMS**

Range of Discrimination indices	Age group 6	Age group 7	Age group 8	Age group 9	Age group 10	Age group 11	Total
Zero or Below	2	3	6	4	2	0	17
1 to 15%	18	15	8	8	10	11	70
16 to 40%	12	14	15	24	28	26	119
41 to 60%	13	16	17	10	10	12	78
61 to 84%	5	2	4	4	0	1	16
Above 84%	0	0	0	0	0	0	0
Total	50	50	50	50	50	50	300

It is evident from the above table that numbers of items that fall out of range 16% to 84% discrimination index are 20, 18, 14, 12, 12 and 11 in tests of age level 6, 7, 8,9,10 and 11 respectively.

**TABLE3: OF ITEMS WITH DIFFICULTY LEVEL < 16**

Age Level	Serial Number of Items
6	Nil
7	Nil
8	15, 27
9	47
10	9, 42
11	Nil

Difficulty level < 16 means that these items are very difficult.

**TABLE4: OF ITEMS WITH DIFFICULTY LEVEL > 84**

Age level	Serial No of Item
6	5,7,10,12,13,14,19,20,26,28,29,30,31,32,33,40,41,43,48,50
7	1,6,14,17,18,19,20,21,22,23,24,25,26,27,29,30,31,35,36,37,38,39,41,42,44,45,47,49
8	1,17,20,30,38,39,41
9	1,2,15,32
10	1,3,19,25,26,28,31,33,34,36,44,45,46,49
11	1,2,12,13,17,20,24,30,31,32,34,35,37,41,45,46,49

Difficulty level > 84 means that these items are very easy because the %age of both high achievers and low achievers is high in these items.

The two properties of an item namely difficulty level and its power to discriminate between good and poor students form the basis for judging its adequacy and worth. Items with extreme values of these two indices were not found acceptable for inclusion in the item pool, because too easy items did not have the capability to discriminate between students and too difficult also tended to be low in discrimination value.

**TABLE5: TEST WISE SERIAL NUMBER OF ITEMS REJECTED WITH DIS:INDEX <=0**

Age level	Serial Number of item
6	1, 29
7	33, 36, 42
8	9,15, 24,27,33 ,40
9	4,7,47,50
10	3,42
11	Nil

The reason for rejection is that these items can't discriminate between low achievers and high achievers, because there is no much difference in the %ages of low and high achievers.

## FINDINGS

Following were the major findings of the study.

1. Most of the items lie in the discrimination index range 16% to 84%.

2. Total number of items in all the tests having discrimination index zero or below zero were 17.
3. 93 items were found with difficulty level above 84%.
4. Only 5 items were found in all tests with difficulty level below 16%.
5. Item No.15 in test of age level 8 is typed incorrectly.
6. Item no.27 in test of age level 8 is missed during typing.
7. Item no.47 in test of age level 9 is typed incorrectly.
8. Item no. 9 in test of age level 10 is either very difficult or not clear.
9. In item.no.42 of test of age level 10 no choice is correct among the given choices.

## **CONCLUSION**

Following were the major conclusions of the study.

1. The items with discrimination index zero or below were rejected. These items were 1, 29 in test of age level six. 33, 36, and 42 in test of age level seven. 9, 15,24,27,33 and 40 in test of age level eight. 4,7,47 and 50 in test of age level nine and 3,42 in test of age level ten.
2. As the total number of items with difficulty level above 84% were 93. This shows that 31% items in the test were easy.
3. As the total number of items with difficulty level below 16% were 5 in all tests. This shows that very difficult items were very few ones.
4. As most of the items fall in discrimination index range 16% to 84%, this shows that these items are good discriminators.

## **RECOMMENDATIONS**

Following recommendation were made in the light of research findings.

1. Intelligence testing requires much more care at every step.
2. The items should be selected with great care; items should be according to the mental level of the students.
3. Items should be performance based if the sample is illiterate and mixed performance and written for literate sample.
4. A big sample should be taken during standardization process.
5. Test should be constructed keeping in view all the important factors of general intelligence.
6. Sample must include at least rural/urban, male/female , poor/rich and if possible black/white peoples.
7. Test administration requires much more care, suitable timing, suitable instructions, and the clarification.
8. Analysis of data should be done with care as it is a work of fatigue and because analysis of big sample requires time and money.
9. The work of intelligence testing should be taken as serious because it is the matter of once future.

## **Acknowledgement**

Taking this opportunity the researchers wish to thank all the direct and indirect supporters who helped, assisted and facilitated the researchers during administration of tests, data collection, its analysis and compilation of results

**BIBLIOGRAPHY**

1. Thorpe, Showick( 2000). General intelligence and test of reasoning. Vikas Publishing house private ltd. , India. Pp34.
2. Sattler, Jerome .M (1981) Assessment of Children intelligence and special abilities. Allyn and Bacon Publishers, America. Pp62.
3. Eysench, H.J (2003). The measurement of intelligence. MTP medical and technical publishing company:ltd, London.
4. Gronlund, Norman E (1998). Preparing criterion reference tests. The mac Millan Company, NewYork London, pp24.
5. Bukhari, Hashmi and Malik.D.M ( 1981). Development and validation of objective type achievement test items. University Grants Commission, Pakistan. Pp35.
6. Atkinson, Hilgard (1995). Introduction to Psychology. Harcourt Brace Publishers, America. Pp399.
7. Eysench, H.J (1976) . Know your own IQ. C. Nicholas and company, Britain.
8. Best. J.W (1959) . Research in Education. Prentice-Hall Publishers, America.pp309.
9. IER Punjab (1978). Test of learning abilities. Printing corporation of Pakistan, Islamabad Pakistan. Pp3-15.
10. Mateen (1996). IQ builder and tester. Maqbool printing press, Pakistan. Pp06.
11. Ebel, R.L and Frisbie 91986). Essentials of Education measurement. Engle wood clif: Prentice-Hall Publishers, America.pp-114.