Bulletin of Education and Research
June 2012, Vol. 34, No. 1 pp.65-76

Development and Validation of an Aptitude Test for Secondary School Mathematics Students

Muhammad Aamir Hashmi, Ashi Zeeshan,
Muhammad Saleem & Rafaqat Ali Akbar

Abstract

The main objective of the study was to develop and validate an aptitude test for secondary school Mathematics students. A multiple choice test of 50 items was finalized after pilot study. The test was administered to a sample of 288 boys and 166 girls of 13 schools of Bahawalpur and Multan district. Separate sheets were provided to students included in the sample. After collecting answer sheets, marks were awarded. Several work sheets were prepared to perform item analysis. Through traditional method of item analysis, difficulty index, discriminatory index and phi coefficient were calculated. The reliability coefficient was found 0.82.

Introduction

Education is the acquisition of knowledge experience. Knowledge is the essence of human life noting enhances man like education. It provides a sound foundation for successful life. By nature, every individual desires to be successful in the domain of his activities. But it has been generally seen that the ratio of being successful has been very low in spite of the fact that people get specific or specialized education and training. The common question arises what is the cause of failure of students on a large scale in the fields of engineering, medical, commerce, law etc. The answer may be the wrong choice of subjects which do not match the potential of the students.

Unfortunately, many pupils do not choose right subjects at school, the right subjects being those which a student can study easily. All too often our choices are determined by other considerations such as how much we like and respect various teachers, how good they are as teachers, what subjects our parents and friends think, are mot "suitable" and of course, what options are offered by the school or college.

Aptitude test plays a vital role to decide the future education (Chatterjee, 2007) of the students. It varies from individual to individual. It indicates individual's and learning ability (Ramsay 2008, Raza 2011), the specific ability needed to facilitate learning a job, suitability, readiness, tendency or natural or acquired disposition or capacity for particular activity (Raza 2011, Reeves 2002), degree of readiness to learn and perform well in a particular situation or in a fixed domain. (Raza 2011, Reed and Wolniak 2005, Cornoet,al., 2002). Salkind and Rashmussen (2007) defined aptitude as a set of characteristics that relate to an individual's ability to acquire knowledge or skills. It further explained that individual differences are related to subsequence learning during a fixed time frame (Salkind and Rasmussen, 2008). Kubiszyn and Borich (2003) defined aptitude as another name for potential or ability. To find the natural ability of an individual in scholastic, vocational or developed component of a competency to do a certain kind of work at a certain level, it needs the aptitude of the student (Raza 2011). Gay (1980) defined the aptitude test as a helper of a teacher to test more realistic expectations of student's abilities and facilitate the identification of under achievers.

Aptitude test measures a student's overall performance across a broad range of mental capabilities. It also includes items which measure more specialized abilities such as verbal and numerical skills-that predicts scholastic performance in educational programs (Oyetunde, 2007). Aptitude tests are cognitive measures used to predict future performance in some activities like school learning (Aiken, 1988; Gronlund, 1981; Sax, 1980; Oyetunde, 2007) The United States Employment service developed the General Aptitude Test Battery (GATB) which is widely used in the United states by state Employment offices and has been made available as a model or starting point for the development of aptitude batteries in other countries (Oystunde, 2007; Tittle 1990). The Armed services vocational Aptitude Battery

(ASVAB) is the most widely used aptitude battery in United State high school. The differential Aptitude Tests (DAT) was designed primarily for educational and vocational counseling in USA secondary schools (Oyetunde, 2007; Tittle 1984).

Mathematics is an important subject in school curriculum. It is more closely related to one's daily life as compared to other subjects. Except one's mother tongue there is no other subject which is more closely related to one's daily life as mathematics. Mathematics is considered to be the father of all sciences. Napoleon remarked that- "*The progress and improvement of mathematics is linked to the prosperity of the state*" (Mahanta and Islam: 1989)

Numerical ability test are designed to measure the candidate's capacity to manipulate are use numbers to correctly solve problems (Ann, 2004; Olatoye and Aderoyba, 2011)

Every individual has specific and varied abilities of different degrees for different professions These abilities include power to think, speak, hear, and see and to understand. It is essential to guide individuals in the right direction at right time in a right way. Success in any field, however, depends first on possessing the appropriate set of basic abilities and then on other factors. But mostly, people follow in father's footsteps or maybe they abide by what parents or friends expect or simply they fell a victim to their aspiration without knowing their peculiar talents and the result is frustration and stress at personal level. There is all too often an unplanned move from school to college and then on to professional qualifications.

It is a sad fact that Pakistan is an under-developed country. And there is a no system of measuring or at least estimating the capacities, potentialities and abilities of students. A larger number of students enter into courses like Medical, Engineering, Commerce, Law, and Agriculture without proper guidance or having no aptitude or potential for the said job and fail to do justice with their choices. Consequently, we suffer from wastage of our best minds, time and money.

If a person's potential or talent could be measured or at least estimated before he begins a course of training much wasted efforts could be spared. All the young ones who aspire to be pilots, Engineers etc., cannot be trained will be ill-afford to do so.

The measurement of aptitude is the crucial one. It is almost a neglected topic on the scenarios of Pakistan. No organized effort has been made in this connection. Owing to expansion and growth of knowledge, it is quite essential to construct tests to measure aptitude of Pakistani students. The present study is an effort to launch a movement in the domain of education.

Background

During the 1920s tests of special aptitude were viewed as measure of specific innate or heredity talents not based on experience. Aptitude tests were first developed by Bingham in 1930s. He refined aptitude as the measure of the probability of the success of an individual with training in a certain type of situation like a job. D. G. Paterson in late 1920s constructed three tests, Minnesota Mechanical Assembly Test, Minnesota Spatial Relation Tests and Minnesota Perp Board. Minnesota spatial test was standardized in 1976. Meier Art tests were approved in 1929 and 1963 respectively. Carl seashore and his colleagues constructed a musical ability test during 1920s and 1930s. Data was jointly constructed by G. K. Benett, H.G. Seashore and A.G. Wesman. This was first published in 1947 and its fourth edition after a number of amendments was published in 1970. After those tests of special aptitude were constructed in almost all domains of Education and training" (Rao, 1990, p259). Aptitude is a trait that varies from individual to individual. That is why speed and accuracy is quite different from one another when they are involved in problem solving. No agreed definition of Aptitude has been advanced by psychologist. It is essentially a natural ability or skill especially in learning. Direct and exact measurement of aptitude is much difficult. Moreover, aptitude and intelligence is not one and the same thing. It is also different from achievement and interest. It is specific and restricted to some extent. According to Woodworth and Marquis (1963),"Aptitude is a predicted achievement." While Aiken (1985: p.224) defined it, "Aptitude has traditionally referred to a person's capacity to profit from further training." Aptitude is an abstract quality or psychological characteristics which cannotmeasure directly. A device used for the purpose of measurement of this potentiality is termed as aptitude test. Different psychologists have described this operational definition. However, it is an agreed fact that predicting future achievement is much more difficult than present achievement. According to Aiken (1985: p.453) an aptitude test is, "A measure of person's ability to profit from further training or experience in an occupation or skill." Woodworth and Marquis (1963: p.26) described it in a different way, "An aptitude test, in itself, is nothing more than an achievement test, but it is used not simply as a measure of present performance in the special task but as an index of future performance in other and broader tasks." Wiersma (1986: p.302), "An aptitude test gives indirect evidence of the existence of the potential." The major object of aptitude testing is to measure varied capacities and potentialities of individuals. It may be difficult in the measurement to separate potential from actual achievement. Aptitude tests in Mathematics are developed to measure the candidates' capacity to manipulate or use numbers to solve problems

correctly (Ann, 2004). These tests imply simple arithmetic ability in an individual. Nunnally (2004) explained, it is the skill to comparatively resolve difficulties in numeral sequencing, make exact mathematical inferences through radical mathematical rational, infer intricate data presented in numerous graphical systems, construe data and draw reasonable inferences. Altogether methods of school studies in various subjects, are also generally discourse kinds of achievement test of which arithmetical aptitude test is one, It can be given straight to contenders or directed as subsets of other tests. Likewise, oral aptitude tests are tests used to infer the capacity of an single to cause with words (Akinboye, 2001). They show the developed competence for conception and communication in an authorized standard like English Language. It may need verbal or inscribed mode of reply as the case can require. Aptitude tests measure the distinctive, developed or advanced factor of capability in facts, indulgent and attitude used in doing certain benevolent of work at a specific glassy (Toplis, 1991). Aptitude may be corporal or psychological; the distinctive nature of aptitude is in distinction to achievement, which represents information that is extended through wisdom (Carr, 2004).

There is some confusion whether an aptitude is innate or inborn or an acquired capacity. For example the scholastic aptitude is highly oriented to the types of abilities learned in formal schooling with an emphasis on verbal and mathematical abilities probably, it is impossible to measure exactly the mental abilities because these are too complex to be measured directly and adequately. It is obvious that achievement tests measure what a people has learned and aptitude tests measure the ability to learn new tasks. According to Wiersma (1986: p. 302)," the development of an aptitude test is a difficult task that requires a good deal of information, effort and measurement exercise. For this reason aptitude tests are seldom self -constructed for a research project. There are numerous aptitude tests available for both general and specific aptitude." Khan (1992, p29) quoted the work done in Pakistan as, "Limited work in this area had been done in different institutions and universities of Pakistan. Mainly the emphasis has been on achievement and intelligence tests. In 1960-61 group scholastic aptitude test for class 10^{th}, 11^{th} and adults was adopted from California test of mental maturity by Amina Talib and Ghulam Hussain, which suit Pakistani culture. In 1972, IER, University of the Punjab, Lahore standardized three scales of scholastic aptitude test for children from class three through class X. Later on in 1984 Naheed Khurshid worked on the Development of Science Aptitude test."

Method

The present Mathematics Aptitude Test was developed after an extensive review of the related literature. The text book of Mathematics for secondary level students was also studied and the concepts were drawn on the base of that book. After getting the concept from the book, the various aptitude tests were studied to construct the items. The only multiple choice items technique were used to construct the items. There were 50 items in all.

Population and Sample

Students of 10^{th} class in all the secondary and higher secondary schools of Multan and Bahawalpur division were selected as the population for the study. Sample for the study was selected on the basis of simple random sampling technique. Schools were selected considering possible assess of the researcher in order to handle the situation properly and effectively. At the first stage schools were selected randomly and then the student of secondary schools by calling their roll numbers randomly (e.g. Roll No. 1, 5, 6, 9, 20, 26 etc.) However, in case of small number of the students in the class, whole of the class was selected for the sample. The target set was thirteen schools including boys' and girls' schools. The number of schools selected from Bahawalpur and Multan division is 6 and 7 respectively of Bahawalpur division 42.86% are girls schools where as 57.14% are boys school. Similarly, the percentage of girls and boys school from Multan district is 66.67% and 33.33% respectively. 71 girls and 202 boys were selected from different schools of Bahawalpur district, 42.25%, 26.67% and 30.98% students were selected from school number 1,2,3, for girls and 10.40%. 34.16% and 45.05% students were selected from boys schools No. 1,2,3 and 4 of Bahawalpur district respectively. On the other hand 95 girls and 86 boys were selected from different schools of Multan district 12.63%, 23.16%, 22.10% and 41.11% of the girls were selected from school no. 1,2,3,4 for girls and 58.14% and 41.86% of the boys were selected from school no. 1 and 2 for boys of Multan district respectively.

Development of the Test

For collection of data, the most useful tool named "test" was used and to determine the content of the test, available literature pertaining to mental abilities and mental assessment was examined. After examining the related literature and model of Bennett, the researcher prepared an aptitude test to test the aptitude of the students of 10^{th} class in the subject of Mathematics. During the construction of the multiple choice items of the test mental level of the students and their texts were seriously

considered. Army aptitude tests were also studied. The researcher initially developed 120 test items under three sections (Arithmetic, Algebra and Geometry). The items were 50, 35 and 35 from Arithmetic's, Algebra and Geometry respectively. There were five options in each item.

Pilot Study and Finalization of the Tool

The test was administered to 60 students including 30 students from each students from each sex of class 10^{th} in various schools of Multan district for pilot study. Relevant information's were collected through answer sheets and a master sheet was prepared in order to tabulate results. One mark for each correct response was awarded and zero for each incorrect response. Negative marking was avoided. Total score was recorded for item analysis. The three major statistical properties were used for item analysis i.e. difficulty index (P) discriminatory index (D) and discrimination power (Q). in the light of the statistical analysis the different changes brought into the test, 40 items were dropped (17 from arithmetic, 11 from Algebra and 12 from geometry). 5 items were replaced (2 from Arithmetic, 1 from Algebra and 2 from geometry). The sequence of 17 items was changed (7 from arithmetic, 4 from Algebra and 6 from geometry). 20 distractors were dropped (9 from arithmetic, 6 from Algebra and 5 from geometry). The sequence of 4 items (3 from arithmetic and 1 from algebra) was changed. 50 items were finalized (20 from arithmetic, 15 from Algebra and 15 from geometry).

Data Collection and Data Analysis

The researcher personally administered the test in different urban and rural schools of Bahawalpur and Multan districts. The instructions given in the test were followed strictly and answer sheets were collected within done time. The students were motivated that test results would not be their results. They acted as an important part of the study quite useful for education. The students were told that there were five options to each item and are name of A, B C, D and E. they were asked to mark the correct answer on the answer sheet only which were separately provided to every students included in the sample. Relevant information were collected through answer sheets and a master sheet was prepared in order to tabulate results. One mark for each response was awarded and zero for each incorrect response. Negative marking was avoided. For scoring an answer key was prepared total scores were recorded for item analysis. The three major statistical properties were used in item analysis i.e. difficulty index (P), Discrimination index (D) and Discrimination power (phi). The

administered test was divided into two halved with respect to even and odd items in the test score of even and odd items were recorded to find the reliability of the test.

Results

The value of difficulty index lies between 0.13 and 0.83. Item number 22 has the largest value (0.83) and item number 23 has the smallest value (0.13). The value of discriminatory index varies from 0.06 to 0.70. Item number has the greatest discrimination power (0.70) and item number 23 has the least discrimination power (0.60). The value of Q coefficient varies from 0.09 to 0.72. Item number 23 has the minimum value (0.09) and item number 28 maximum value (0.72). The reliability coefficient was calculated which is 0.82.

Table 1

Comparative study of P, D and Q of the test

Item No.	P	D	Phi
1	0.69	0.21	0.22
2	0.82	00.32	0.38
3	0.75	0.35	0.39
4	0.69	0.34	0.47
5	0.60	0.44	0.44
6	0.50	0.26	0.26
7	0.54	0.51	0.52
8	0.51	0.70	0.63
9	0.37	0.34	0.35
10	0.66	0.63	0.66
11	0.75	0.31	0.35
12	0.75	0.38	0.42
13	0.82	0.31	0.39
14	0.40	0.35	0.35
15	0.65	0.53	0.54
16	0.66	0.35	0.36
17	0.45	0.34	0.34
18	0.59	0.55	0.56
19	0.59	0.35	0.36
20	0.52	0.57	0.57
21	0.51	0.58	0.58

22	0.83	0.34	0.42
23	0.13	0.06	0.09
24	0.74	0.44	0.44
25	0.37	0.19	0.20
26	0.26	0.31	0.32
27	0.28	0.13	0.15
28	0.72	0.34	0.72
29	0.58	0.40	0.41
30	0.38	0.42	0.43
31	0.45	0.42	0.24
32	0.29	0.30	0.31
33	0.40	0.35	0.35
34	0.52	0.47	0.43
35	0.46	0.62	0.61
36	0.20	0.17	0.26
37	0.46	0.41	0.41
38	0.46	0.54	0.52
39	0.51	0.62	0.62
40	0.42	0.38	0.39
41	0.39	0.42	0.43
42	0.37	0.26	0.27
43	0.42	0.48	0.47
44	0.42	0.29	0.29
45	0.36	0.22	0.24
46	0.44	0.57	0.61
47	0.26	0.18	0.19
48	0.60	0.22	0.23
49	0.62	0.16	0.17
50	0.44	0.22	0.22

Discussion and Recommendations

Average value of the “P” of the test is 0.52 which represents that the test approximate to an ideal test having twenty one items with difficulty index (P) greater than the average value of “P” and twenty six items with difficulty index (P) smaller than the average value of “P”. it reflects that the test items are capable to differentiate the aptitude of the students to a greater extent. According to the value of “P” item number 2, 13 and 22 can be treated as the easiest item. On the other hand, items no.

23, 27, 32, 36, and 47 are difficult items with respect to the sample. On the basis of the results of the values of "D" it can be claimed that there is not a single item showing negative value with respect to the sample selected. Item no. 23 is the only item having the value less than 0.10 which demands a major change in the item or it may be replaced. The item no. 25, 27, 36, 47 and 49 demand a minor change in the item. The value of discriminatory power (Q) reflects that there is not a single item having zero or negative values. Item number 23 is the only item with Q value less than 0.10. Item number 27, 47 and 49 have Q values less than 0.20 and demand minor changes in their construction.On the basis of the item analysis down with the help of "P" "D" and "Q", it is concluded that item number 23 is most difficult item of the test and it may be rejected.

Further, item number 27 and 47 require change in their construction. However, item number 32, 36, 44 and 49 need modifications to some extent. Item number 22 also demand some changes in the construction. The drop out of item number 23 and necessary changes in the item number 22, 27, 32, 36, 44, 47 and 49 are essential to enhance the reliability of the test. For the standardization of the test sample size should be increased. More institutions of other district should be included in the sample. More schools from rural areas should be included in the sample to make it more representative. Efforts should be made to collect information about the personal bio-data of the students e.g. date of birth, parent's education, parent's profession, favorite subject, favorite teacher and the subject of favorite teacher. Teachers should make use of aptitude test in order to measure students' aptitude in the subject of mathematics.

References

Aiken, L. R. (1985) "*Psychological Testing and Assessment"*, fifth edition, Boston: Allyn and Bacon Inc

Aiken, L. R. (1988). Psychological testing and assessment (6th ed.). Boston:Allyn and Bacon, Inc.

Akinboye, J.O. (2001). Concepts in FormalEducation, Ibadan: University Press Limited.

Ann, P. S. (2004). Measurement, Assessment andEvaluation, Lagos: Concepts Publications Limited.

Carr, H.O. (2004). Education Psychology, New York:McGraw Hill Ltd.

Chatterjee, A. (2007).A Controlled and Effective Education System for a country. Retrieved on February 24, 2008 from http://www.indianstudentresearch.blogspot.com/2007/11/controlled-andeffectiveeduca tion.html

Gronlund, N. E. (1981). Measurement and evaluation in teaching. New York:Macmillan Publishing Co.

Hashmi, M. A. (2000), *"Standardization of An Intelligence Test for Middle Level Students"* Unpublished Thesis, Department of Education Bahauddin Zakariya University Multan

Khan, M. N. T. (1992), *"Semi-Standardization of An Intelligence Test for The Students of Class VI"*, Unpublished Thesis, Department of Education BahauddinZakariya University Multan

Nunnally, J.O. (2004). Memory, New York:Macmillan Press.

Ramsay, T. (2008).Advantages of Multiple Choice Aptitude Tests. Retrieved on May 25, 2008 from http://www.ramsaycorp .com/categories/aptitude.asp

Raza, M. A., and Shah, F. A. (2007) , " Impact of Favorite Subject Towards the Scientific Aptitude of the Students at Elementary Level" Pakistan Journal of Social Sciences (PJSS) Vol. 31 No.1 (June 2011) pp 135-143

Reed, E. J. and Wolniak, G. C. (2005).Concept of Aptitude., Retrieved on February 23, 2007 from http://www.sociology.org/content/2005/tier1/reed_wolniak.pdf

Reeves, J. F. (2002).Aptitude Assessment. Retrieved on January 15, 2005 from http://www.theworksuite.com/id15.html

Rao, S. N. (1990) ,*"Educational Psychology"*, New Delhi: Wiley Eastern Ltd.

Saleem, M. (1994),*"Development of An Aptitude Test of Mathematics at SSC Level"* Unpublished Thesis, Department of Education Bahauddin Zakariya University Multan

Toplis, K. (1991). Behavioural ModificationTechniques, New York: Prentice Hall Publishers.

Wiersma, W. (1986), *"Research Methods in Education: An Introduction",* (4th edition) Boston: Allyn and Bacon Inc

Woodworth, R. S. and Marquis, D. G. (1963) , *"Psychology"* 20th Edition, Strand: Methuen and Co. Ltd